Measurement of s-channel single top-quark production with the ATLAS detector using total event likelihoods

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Physik
eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Patrick Rieck

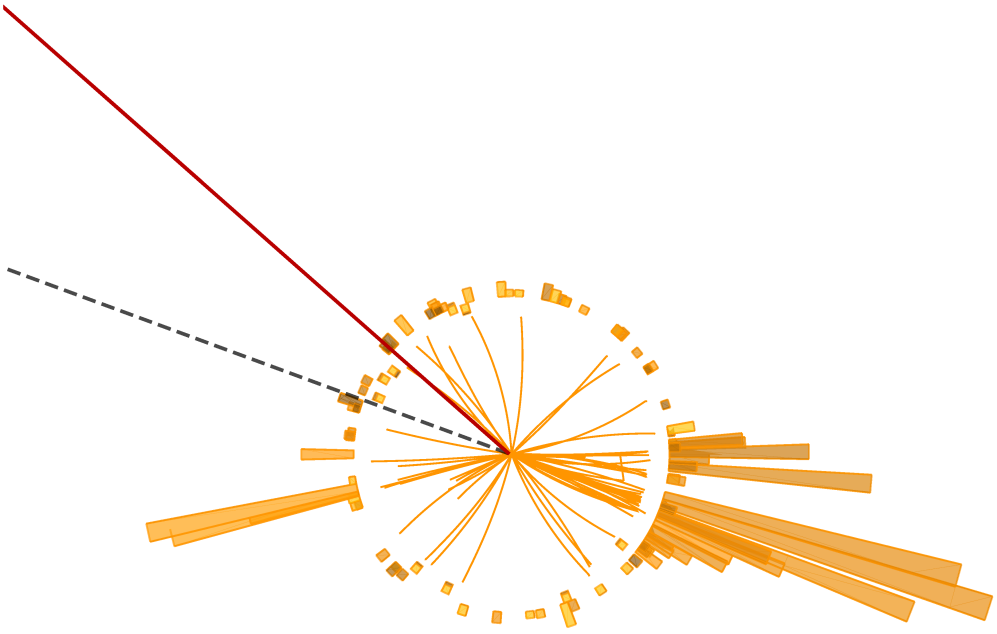
Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Thomas Lohse
2. Prof. Dr. Heiko Lacker
3. Prof. Dr. Ulrich Husemann

eingereicht am: 20. Oktober 2015
Tag der mündlichen Prüfung: 5. April 2016

Measurement of s-channel single top-quark production with the ATLAS detector using total event likelihoods



Dissertation

Patrick Rieck

Abstract

A measurement of s-channel single top-quark production in proton-proton collisions at a centre-of-mass energy of 8 TeV is presented. The data set has been recorded with the ATLAS detector at the LHC and corresponds to an integrated luminosity of $20.3\,\mathrm{fb}^{-1}$. Collision events are selected so that a subset of the data is obtained where the signal fraction is relatively high. Selected events contain one isolated electron or muon, missing transverse momentum and 2 jets, both of which are induced by b-quarks. All of these objects have large transverse momenta. The resulting set of events is still dominated by background processes, most notably top-quark pair production and the production of W bosons in association with jets. In order to further separate the signal from the backgrounds, several approximate event likelihoods are computed. They are based on different hypotheses regarding the scattering process at hand. Together they result in a function of the measured momenta which allows for the desired separation of the signal process. A statistical model of the corresponding distribution is used in a fit to the measured data. The fit results in a signal significance of 3.4 standard deviations and a total cross section of $\sigma_{s-\mathrm{channel}}^{\mathrm{LHC},\,8\,\mathrm{TeV}} = 5.3^{+1.8}_{-1.6}\,\mathrm{pb}$. This is the first evidence for s-channel single top-quark production in proton-proton collisions. The results agree with the standard model prediction.

Zusammenfassung

Es wird eine Messung der s-Kanal Einzel Top-Quark Produktion in Proton-Proton Kollisionen bei einer Schwerpunktsenergie von 8 TeV vorgestellt. Der verwendete Datensatz wurde mit dem ATLAS Detektor am LHC aufgenommen und entspricht einer integrierten Luminosität von $20.3\,\mathrm{fb}^{-1}$. Kollisionsereignisse werden selektiert, sodass der resultierende Anteil der Signalereignisse relativ hoch ist. Selektierte Ereignisse enthalten ein isoliertes Elektron oder Myon, fehlenden Transversalimpuls und zwei Jets, die durch b-Quarks induziert wurden. Alle Objekte haben hohe transversalimpulse. Auch nach dieser Selektion überwiegen Untergrundprozesse, insbesondere die Paarproduktion von Top-Quarks und die Produktion von W-Bosonen begleitet von Jets. Um den Signalprozess weiter von den Untergründen zu trennen, werden mehrere Wahrscheinlichkeitsdichten näherungsweise berechnet. Sie unterscheiden sich hinsichtlich der Annahme des zugrundeliegenden Streuprozesses. Zusammen ergeben sie eine Funktion der gemessenen Impulse, mit deren Hilfe das Signal weiter von den Untergründen getrennt werden kann. Ein statistisches Modell der entsprechenden Verteilung wird an die Messdaten angepasst. Diese Messung ergibt eine Signifikanz des Signalprozesses von 3.4 Standardabweichungen und einen totalen Wirkungsquerschnitt von $\sigma_{s-\mathrm{Kanal}}^{\mathrm{LHC},\,8\,\mathrm{TeV}} = 5.3^{+1.8}_{-1.6}\,\mathrm{pb}$. Dies ist die erste signifikante Messung der s-Kanal Einzel Top-Quark Produktion in Proton-Proton Kollisionen. Die Ergebnisse stimmen mit der Vorhersage des Standardmodells überein.

Contents

Introduction **1**

1. Theoretical foundations **5**
1.1. The standard model of particle physics 5
1.1.1. Basic principles 6
1.1.2. Model of elementary particles and interactions 10
1.2. Hadronic collision phenomenology 18
1.2.1. The factorization theorem 18
1.2.2. Jets of strongly interacting particles 21
1.2.3. Event rates at hadron colliders 23
1.3. Top-quark phenomenology 24
1.3.1. Top-quark decay 26
1.3.2. Top-quark production in hadronic collisions 27
1.3.3. CKM matrix elements 33
1.3.4. Single top-quark cross section predictions 35
1.3.5. Searches for s-channel single top-quark production 39
1.4. Simulation of hadronic collision events 42
1.4.1. Factorization of scattering processes 42
1.4.2. Selected event generators 46

2. Experimental setup **49**
2.1. The Large Hadron Collider 49
2.1.1. CERN's accelerator complex 50
2.1.2. Magnets 52
2.1.3. Proton beams 52
2.2. The ATLAS detector 54
2.2.1. Magnets 58
2.2.2. Inner tracking detectors 58
2.2.3. Calorimeters 61

2.2.4. Muon spectrometer 66
2.2.5. Electronics, trigger and data acquisition 69
2.3. LHC and ATLAS performance in 2012 72
2.4. ATLAS event simulation 74
2.5. Physics object reconstruction 77
2.5.1. Primary vertices 77
2.5.2. Electrons 78
2.5.3. Muons 83
2.5.4. Jets 88
2.5.5. Missing transverse momentum 97

3. Single top-quark measurement **99**
3.1. Analysis basics 99
3.1.1. Data set 100
3.1.2. Modelling of processes 101
3.1.3. Event selection 109
3.1.4. Event yields and control distributions 115
3.2. The method of total event likelihoods 121
3.2.1. Ansatz 123
3.2.2. Implementation 130
3.2.3. Scattering processes 135
3.2.4. Transfer functions 137
3.2.5. Likelihood control distributions 143
3.2.6. Event classification 147
3.2.7. Further studies 149
3.3. Systematic uncertainties 156
3.3.1. Sources of uncertainties 156
3.3.2. Impact on rates 162
3.3.3. Impact on discriminant shapes 162
3.4. Search for s-channel single top-quark production 166
3.4.1. Profile likelihood fit 168
3.4.2. Signal significance 174
3.4.3. Cross section confidence interval 178
3.5. Future prospects 183

Conclusion **189**

Appendix **191**

A. Simulation samples **193**

B. Single top-quark cross section computations **199**
B.1. Partonic cross sections at NLO in QCD 200
B.2. Renormalization and factorization scale dependence 203
B.3. Hadronic cross section results 205

C. Additional studies **209**
C.1. Standard model fit of control distributions 209
C.2. Di-lepton veto cut 215
C.3. Resolution functions used for TELLs 217
C.4. Binning and modelling of the TELL discriminant 220
C.5. W+jets modelling study 223
C.6. Fit results of the measurement 227

D. Tell phase space generation **231**
D.1. Generic phase space 232
D.2. W+jets production 233
D.3. Single top-quark production 235
D.4. Top-quark pair production 240

References **247**

List of Figures **265**

List of Tables **271**

Introduction

On July 4th 2012, the ATLAS and CMS collaborations at CERN's Large Hadron Collider announced the discovery of a particle whose properties correspond to those of the standard model Higgs boson [1, 2]. This achievement marks the completion of the search for the elementary particles predicted by the standard model (SM). Now this model can be summarised as a distinct description of elementary particles and their interactions, while the vast majority of its predictions to which experiments up to today have been sensitive are confirmed. In particular, it describes all known visible and stable particles, including bound states such as nuclei and atoms. Furthermore, it describes particles which are heavier than the stable ones. They can be classified in similar patterns known as the second and third generation. Interactions are derived from basic symmetry principles which give rise to bosons coupling to the fermionic matter particles. They include all known forces except for gravity. Finally, the particle property of mass is described in the context of the Higgs mechanism.

Nevertheless, the end of searches for SM particles is by no means the end of particle physics. Despite its predictive power, the SM fails to answer major questions of physics. It does not explain the very existence of matter in the universe and the fact that matter was annihilated by antimatter in the early universe only to an extend which still allows for the structure of the universe which we see. Admittedly it predicts CP-violation which is a premise for the baryogenesis, but it does not do so to a sufficient amount. It does not explain the existence of dark matter, which is necessary to explain the larger scale structure of the universe and cannot be directly detected by optical means. It also does not predict the way in which neutrinos obtain mass and it does not, as indicated above, describe the gravitational interaction of particles or at least why gravity is much weaker than all other known forces. In addition to these strict deficiencies, a weakness of the SM is the fact that it does not fix many of its parameters but instead requires them to be chosen according to experimental results. For these reasons the SM can be assumed to be an effective description of a more fundamental model, similar to the periodic table of the elements described by chemistry before the rise of quantum mechanics.

To find physics beyond the SM is the task of the Large Hadron Collider (LHC). In fact, this is its only remaining task now that a candidate for the Higgs boson has been discovered. In its proton-proton collisions with energies and rates higher than ever reached before in artificial particle accelerators such new phenomena could show up. In order to record them, several detectors have been built around dedicated interaction points. ATLAS [3] and CMS [4] are multi-purpose experiments mainly designed for searches at the energy frontier, while ALICE [5] and LHCb [6] are more specialized detectors. In order to find phenomena beyond the SM, these experiments, in particular ATLAS and CMS, perform a broad range of tasks. Apart from direct searches for new phenomena it is for several reasons also mandatory to thoroughly investigate the phenomena described by the SM. First of all, almost all of the assumed scenarios beyond the SM compete with large SM backgrounds. Being a hadron collider, the LHC delivers complex collision events. Therefore a good understanding of SM process is a prerequisite for searches for new phenomena. Secondly, the character of new phenomena is of course unknown. They could show up as minor deviations from the SM expectation, e.g. in terms of properties of known particles. Last but not least, SM measurements are laboratories which allow one to develop new techniques which improve the understanding of the complex collision events. Once these techniques are established, they often can be translated to other tasks including searches for new phenomena.

For all of these reasons this thesis is dedicated to the measurement of the s-channel single top-quark production cross section. The top-quark is the heaviest of all elementary particles known today. With its mass of about 173.3 GeV [7] it is approximately as heavy as a whole tungsten atom. In consequence, its decay phase space is so large that it decays before it would hadronize which makes it an ideal probe for studies of quark couplings. The decay chain of top-quarks leads to complex signatures. In particular they involve large missing transverse momentum which makes top-quark production an important background for many searches for new phenomena. The basic top-quark production modes at the LHC are the dominant pair-production mediated by strong interactions and single top-quark production mediated by weak interactions.

Single top-quark production offers exceptional opportunities to measure CKM matrix elements related to the top-quark, which in turn allows one to test the SM. It is a unique source of highly polarized top-quarks which can be used to investigate couplings of this particle in detail. Furthermore, several scenarios of phenomena beyond the SM affect single top-quark production, for instance heavy charged bosons or sizeable flavour changing neutral currents. Single top-quark production is usually distinguished

by the four-vector of the exchanged W boson being spacelike (t-channel) or timelike (s-channel). This distinction makes sense as the interference of the two contributions is small due to their different colour structure. Among these processes, s-channel single top-quark production has by far the smallest cross section at the LHC according to the SM. Therefore advanced methods of signal discrimination are developed for the search for this process presented in this thesis.

The data set used for this search consists of proton-proton collisions at a centre-of-mass energy of 8TeV which have been recorded with the ATLAS detector in 2012. The integrated luminosity amounts to $L = 20.3\,\mathrm{fb}^{-1}$. The key point of this analysis is the technique developed and applied to separate the s-channel signal from its much more frequent backgrounds. This technique is know as the matrix element method or, more specific, as the method of *t*otal *e*vent *l*ike*l*ihoods (TELL).

The thesis is structured as follows: in chapter 1 the SM and the role of the top-quark within this model are summarised. Emphasis is put on the cross section predictions to which the experimental analysis refers. Chapter 2 describes the employed experimental setup. Finally, chapter 3 presents the search for s-channel single top-quark production.

This thesis is a result of common efforts. Working in cooperation with the ATLAS collaboration, in particular including the experimental particle physics research group at Humboldt-Universität zu Berlin, the author is responsible for major aspects of the measurement presented in chapter 3. In particular, the method of total event likelihoods is developed at Humboldt-Universität. In addition, studies of single top-quark cross section predictions presented in this thesis are performed by the author using a dedicated, efficient computer program. It makes use of pre-existing next-to-leading order cross sections and the author contributed to their incorporation.

Chapter 1.

Theoretical foundations

"Le savant n'étudie pas la nature parce que cela est utile;
The scientist does not study nature because it is of use to do so;
il l'étudie parce qu'il y prend plaisir
he does so because it gives him pleasure
et il y prend plaisir parce qu'elle est belle.
and it gives him pleasure because nature is beautiful."

— Henri Poincaré, *Science et Méthode*

1.1. The standard model of particle physics

The SM is indeed, as the name suggests, a *model*. The more fundamental *theories* which it is based on are special relativity and quantum mechanics. They are combined in the context of relativistic quantum field theory (QFT). In the present section, the theoretical basis of the SM, given by a couple of principles, is introduced first. Then the model itself is presented. Throughout this thesis, actions and velocities are measured in units of the fundamental constants of the underlying theories, $\hbar = 1$ and $c = 1$.

1.1.1. Basic principles

Relativistic invariance

One of the foundations of elementary particle physics is the Poincaré group, which describes transformations of spacetime $x \to x'$, namely translations and Lorentz transformations:

$$x = (t, \mathbf{x}) = (x^0, x^1, x^2, x^3)\,, \quad x^\mu \to x'^\mu = a^\mu + \Lambda^\mu{}_\nu x^\nu\ . \tag{1.1}$$

Here a is a translation vector, while Λ is a Lorentz transformation, $\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma g_{\mu\nu} = g_{\rho\sigma}$ with g being the Minkowski metric. Here and in the following, repeated indices are summed over. Models of elementary particles need to be invariant under translations and under Lorentz transformations which are connected to the identity. On the other hand, time reversal, $(t, \mathbf{x}) \to (-t, \mathbf{x})$, and parity, $(t, \mathbf{x}) \to (t, -\mathbf{x})$, are not required to be valid symmetries since experimental evidence has ruled them out.

In quantum mechanics, symmetries can be represented by unitary operators defined on a Hilbert space. Accordingly, each spacetime transformation is represented by such an operator. They are in turn related to hermitian operators which generate the transformations. Aside from boost generators, these are the energy or Hamiltonian H, the linear momentum $\mathbf{P}$ and the angular momentum $\mathbf{J}$, all of which are conserved.

As in every quantum theory, it is useful to find a complete set of commuting observables in order to distinguish all states by their corresponding eigenvalues. Types of elementary particles correspond to states which share the same, invariant eigenvalues, or spacetime quantum numbers. [1] More precisely, they are unitary, irreducible representations of the Poincaré group. The invariants of the Poincaré group are the mass m, given by $m^2 = H^2 - \mathbf{P}^2$, and the spin S, given by the angular momentum $\mathbf{J}^2 = S(S+1)$ in a particles rest frame in case of massive particles. While the latter can occupy $2S+1$ spin states, massless particles can only occupy two helicity states $\mathbf{S} \cdot \mathbf{P}/|\mathbf{P}| = \pm S$.

Quantum fields

In ordinary quantum mechanics, wave functions describe distributions of a fixed number of particles in space. While this approach is highly successful in case of low energies,

[1] Other types of quantum numbers, called charges, matter too. They are discussed below.

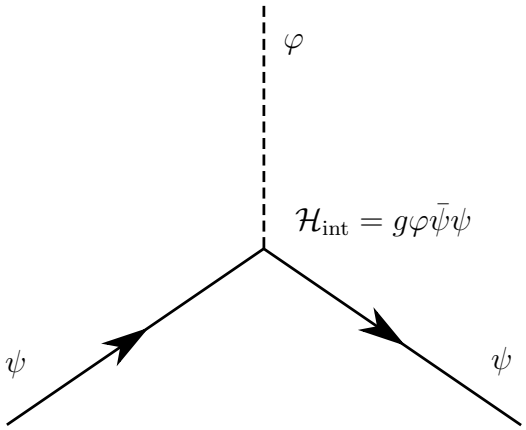


Figure 1.1.: Yukawa interaction as a Feynman diagram. The interaction term $g\bar{\psi}\psi\varphi$ of Eq. 1.2 describes the scattering of fermions ψ and bosons φ with a coupling strength g. Its leading order contribution is given by one annihilation and one creation of the fermion, while a boson is being exchanged with a scattering partner not shown here.

several inconsistencies arise from wave equations if kinetic energies are on the order of the involved particle masses or greater. Indeed, this approach is not expected to work out. At high energy scales, creations and annihilations of particles take place due to Einstein's relation $E = mc^2$. But in the ordinary approach, there is no mechanism involved which allows for such processes to occur. In QFT, on the other hand, operators are introduced which are functions of spacetime. They describe the creation and annihilation of particles, thereby mapping different multi-particle states onto each other, including the ground state, or vacuum. Once anti-particles are included in this approach, QFT can overcome inconsistencies encountered in ordinary quantum mechanics.

One approach to a QFT is to consider first of all a classical field theory, say with a Lagrangian density $\mathcal{L}(\varphi(x), \partial_\mu\varphi(x))$ which depends on a real field $\varphi(x)$. Then φ is declared to be a quantum field and commutation relations among φ and $\pi = \partial\mathcal{L}/\partial\dot{\varphi}$ are introduced by analogy to the case of coordinates in the configuration space of ordinary quantum mechanics. Accordingly, $\mathcal{L}$ becomes an operator, and so does the Hamiltonian $H = \int \mathrm{d}^3x\, \mathcal{H}$ with $\mathcal{H} = \pi\dot{\varphi} - \mathcal{L}$. Using the time-evolution operator $U(t) = e^{-iHt}$, transition amplitudes between incoming and outgoing states of scattering processes can be computed and compared with experimental results. A simple example of a QFT is known as Yukawa-theory,

$$\mathcal{L}_{\text{Yukawa}} = \underbrace{\frac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) - \frac{1}{2}m^2\varphi^2}_{\text{Klein-Gordon}} + \underbrace{\bar{\psi}\left(i\partial\!\!\!/ - m\psi\right)\psi}_{\text{Dirac}} - \underbrace{g\bar{\psi}\psi\varphi}_{\text{Yukawa}}\,, \tag{1.2}$$

where $\not{\partial} = \gamma^\mu \partial_\mu$ with gamma matrices γ^μ, $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, and $\bar{\psi} = \psi^\dagger \gamma^0$. In the classical field theory, the first two terms yield the Klein-Gordon and the Dirac equation, respectively, by means of the principle of least action. As QFTs, they describe the unperturbed propagation of the respective field quanta. The term $g\bar{\psi}(x)\psi(x)\varphi(x)$ describes interactions of the fermion ψ with the boson φ at the spacetime point x with a coupling strength g. Concerning the computation of observables, the most common and most effective approach is to make use of perturbative expansions in the couplings, as sketched by the Feynman diagram of Fig. 1.1.

Renormalizability

Once a QFT is evaluated beyond the leading order of perturbation theory, ultraviolet divergences arise. They reveal the fact that the bare fields and couplings which appear in a Lagrangian involve singularities. Renormalization allows one to interpret models in a meaningful way despite the appearance of these divergences. It is based on the fact that QFTs do not predict the absolute strength of interactions, since the coupling constants are not fixed anyway. However, QFTs should predict the relative strengths of couplings given different interaction scenarios. Once amplitudes are related among each other by means of a reference, or renormalization scale, they are finite and correspond to observables.

The procedure of renormalization cannot be applied to every model. In particular, interaction terms with couplings of negative mass dimension, e.g. $\lambda\varphi^n$ with $n > 4$, are excluded by the requirement of renormalizability. [2] If such terms would be accepted as part of a model, higher order corrections were ill-defined and could only be handled ad hoc by means of high energy cutoffs. Hence the requirement of renormalizability is important for a model to have predictive power and it drastically decreases the number of possible interaction terms. In fact, the SM contains all renormalizable interaction terms of fields of spin $S = 0$, $1/2$ and 1 [8–11].

Gauge invariance

In order to describe the forces of nature, additional fields must be added to the Lagrangian of Eq. 1.2. In the SM, forces couple to charges carried by particles and these charges are conserved. By means of Noether's theorem, every conservation law is connected

[2] For the field φ which enters $\mathcal{L}_{\mathrm{Klein-Gordon}}$ it follows from $[\int \mathrm{d}^4x\mathcal{L}] = 0$ and $[\partial_\mu] = 1$ that $[\varphi] = 1$.

to an invariance, or symmetry of the model. In fact, gauge symmetries based on the special unitary groups $\mathrm{SU}(N) = \{U \in \mathbb{C}^{N\times N} \mid UU^\dagger = 1 \;\&\; \det U = 1\}$ define models known as Yang-Mills theories, which form the basis of the SM [12]. Their elements can be parametrised in terms of real-valued functions of spacetime $\alpha^a(x)$ and hermitian operators T^a,

$$U(x) = 1 + i\alpha^a(x)\, T^a + \mathcal{O}((\alpha^a)^2) \ . \tag{1.3}$$

The $N^2 - 1$ independent generators obey the commutation relations $[T^a, T^b] = if^{abc}T^c$, where f^{abc} are numbers. Starting with the free theory, the derivative ∂ is replaced by the so-called covariant derivative D,

$$D_\mu = \partial_\mu - igA^a_\mu T^a \ , \tag{1.4}$$

which introduces $N^2 - 1$ gauge fields, or forces A^a_μ. The extracted constant g is arbitrary at this point but turns out to be convenient. The gauge transformations of the fields are

$$\psi \to (1 + i\alpha^a T^a)\psi + \mathcal{O}\left(\alpha^2\right) \ , \tag{1.5}$$

$$A^a_\mu \to A^a_\mu + \frac{1}{g}\partial_\mu \alpha^a + f^{abc}A^b_\mu \alpha^c + \mathcal{O}\left(\alpha^2\right) \ . \tag{1.6}$$

The kinetic term of the gauge fields involves the field strength tensor $F^a_{\mu\nu}$,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \ . \tag{1.7}$$

Finally, the Yang-Mills Lagrangian, invariant under $\mathrm{SU}(N)$ gauge transformations, is

$$\mathcal{L}_{\mathrm{Yang-Mills}} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \ . \tag{1.8}$$

Concerning the creation of models based on these symmetries, two questions remain to be answered. Which one is the gauge group adequate for the description of nature, if any, and what does the fermion space look like or, equivalently, what are the interior quantum numbers of the fermions. A very successful answer to these questions, namely the SM, is presented in the next section.

1.1.2. Model of elementary particles and interactions

The SM of particle physics is based on a product of gauge groups where strong interactions on the one hand, and electroweak interactions on the other hand factorize. The masses of particles emerge due to the so-called Higgs mechanism. In the following, the particular sectors of the SM are presented. Mostly the lines of thought presented in Ref. [13] are followed. As the first part, a kinetic term is needed for every type of fermion,

$$\mathcal{L}^{\text{free}}_{\text{fermions}} = \sum_f \bar{\psi}_f i \partial\!\!\!/ \psi_f \ . \tag{1.9}$$

There are three generations of fermions, each of which consists of two quarks, one charged lepton and one neutrino. Further concretisations of the fermion sum, related to interior quantum numbers, are given in the course of this section.

Strong interactions

Quarks and gluons are involved in the strong interactions. This force must be rather special in order to enforce the confinement of quarks in bound states, called hadrons. It turns out that $\mathrm{SU}(3)_C$ is an appropriate choice, where C stands for colour. Each quark comes in a so-called colour triplet,

$$\Psi_q = \begin{pmatrix} \psi_q^r \\ \psi_q^g \\ \psi_q^b \end{pmatrix} \ , \tag{1.10}$$

where r, g, g label states of the fundamental representation of the gauge group, while q labels the quark flavour. The interactions of this model, called Quantum Chromodynamics (QCD), follow from the Yang-Mills Lagrangian of Eq. 1.8,

$$\mathcal{L}^{\text{int}}_{\text{strong}} = g_S G^a_\mu j^{a}_{G,\,q}{}^{\mu} \ , \quad j^{a}_{G,\,q}{}^{\mu} = \bar{\Psi}_q \gamma^\mu T^a \Psi_q \ , \tag{1.11}$$

$$\mathcal{L}^{\text{kinetic}}_{\text{gluon}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} \ . \tag{1.12}$$

G^a_μ and $G^a_{\mu\nu}$ label the gauge fields, or gluons, and their field strengths, respectively, while g_S denotes the strong coupling strength. In principle, QCD could be extended by a term violating the CP symmetry, which combines charge conjugation and parity, but

experimentally no evidence for CP violation in strong interactions has been found to date. Hence, the corresponding term is neglected.

Electroweak interactions

In weak interactions, there are positively and negatively charged currents as well as neutral currents. In total this suggests the presence of three weak gauge bosons. The SU(2) group comes with the appropriate number of generators. In order to incorporate the photon, an additional U(1) symmetry is imposed, $\mathrm{U}(1) = \{U \in \mathbb{C} \,|\, UU^* = 1\}$. However, its generator cannot correspond to the electric charge Q directly, because in this case weak and electromagnetic interactions would factorize and hence charged currents would not exist. Instead, the U(1) generator is called weak hypercharge Y, while Q must be a linear combination of SU(2) generators and Y [14].

A main issue of the electroweak sector is the fact that weak interactions are indeed much weaker than electromagnetism. This difference can be explained if the weak gauge bosons carry mass [15, 16]. In this case, weak interactions are suppressed at energy scales different from the boson mass scale. However, corresponding mass terms like $\frac{1}{2}m_A^2 A_\mu A^\mu$ are not part of the Yang-Mills Lagrangian (Eq. 1.8) and violate gauge invariance. Still, masses can be generated by spontaneous symmetry breaking which means that the ground state has a lower symmetry than the overall model. For this purpose, an SU(2) doublet Φ and an appropriate potential are introduced,

$$\Phi = \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix} , \tag{1.13}$$

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 . \tag{1.14}$$

Figure 1.2 shows an illustration of this potential for a single complex scalar field $\varphi_1 + i\varphi_2$. It is symmetric under rotations in the plane of the two real-valued fields φ_1 and φ_2. Accordingly, there is an infinite multitude of ground states. For any fixed ground state, two degrees of freedom can be distinguished. Variations along the radial direction are excitations and correspond to a massive boson. Variations along the azimuthal direction, on the other hand, constitute an equipotential degree of freedom. In the case of the SU(2) doublet potential of Eq. 1.14, there are three such symmetry directions, corresponding to the three generators of the group. This fact can be used, for instance, to make all fields φ_i

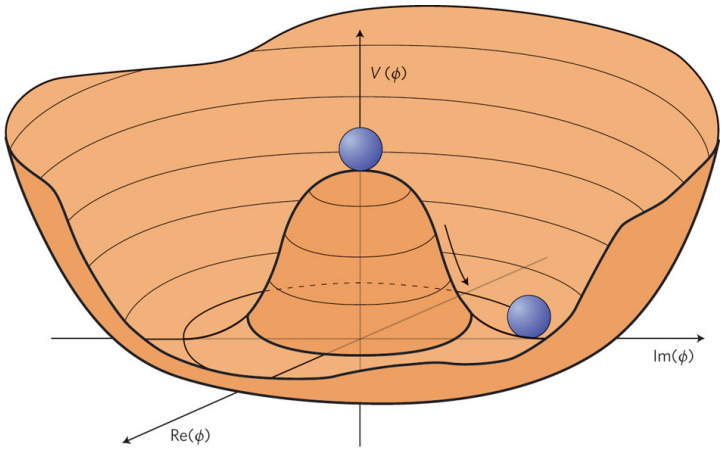


Figure 1.2.: Higgs Potential in case of a single complex scalar field. The ground state, illustrated by the ball sitting in the potential valley, has a lower degree of symmetry than the overall potential. Radial variations change the potential, while azimuthal ones do not. In the SM, this spontaneous breaking of symmetry gives rise to gauge boson masses, related to the equipotential variations, and to the Higgs boson, related to radial variations [23].

except for φ_3 vanish at every point in space time by means of local gauge transformations. In particular, the ground state in this case is

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad , \quad v = \sqrt{\frac{\mu^2}{\lambda}} \, . \tag{1.15}$$

Once the gauge is fixed in this way, the fields φ_1, φ_2 and φ_4 do not appear explicitly anymore. Instead, the gauge bosons associated with the symmetries affected by gauge fixing acquire longitudinal polarisation states as additional degrees of freedom. They become massive. This way of introducing masses is known as the Higgs-mechanism [17–22].

Before these relations can be evaluated explicitly, a weak hypercharge eigenvalue must be assigned to the Higgs doublet. With $Y_\Phi = 1/2$ its gauge transformation is

$$\Phi \to \Phi' = (1 + i\alpha^a T^a)(1 + i\frac{1}{2}\alpha^Y)\Phi \, . \tag{1.16}$$

If for instance the Pauli matrices are used as the SU(2) generators, in particular $T^3 = \mathrm{diag}\,(1/2, -1/2)$, the ground state Φ_0 of Eq. 1.15 remains invariant under local gauge transformations with $\alpha^3 = \alpha^Y$. Accordingly, one gauge boson remains massless. It

must be the photon and hence the generator of the electric charge is

$$Q = T^3 + Y \,. \tag{1.17}$$

The masses of gauge bosons appear in the following part of the Lagrangian,

$$\begin{aligned} \mathcal{L}_{W,Z}^{\text{mass}} &= (D_\mu \Phi_0)^\dagger \, (D^\mu \Phi_0) \\ &= m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2} m_Z^2 Z_\mu^0 Z^{0\mu} \;, \end{aligned} \tag{1.18}$$

The fields $W^\pm$ and Z^0 denote linear combinations of the SU(2) and U(1) gauge fields. They are massive and their masses are proportional to the vacuum expectation value of the Higgs field, v. Another linearly independent field is the photon A_μ, which remains massless. The covariant derivative D_μ involves a term $-ig \sin\theta_w A_\mu Q$ with $\sin\theta_w = g'/\sqrt{g^2 + g'^2}$, where g and g' are the coupling strengths of the SU(2) and U(1) gauge groups, respectively. Hence the electric charge is $e = g \sin\theta_w$.

The kinetic terms of the gauge bosons introduce various couplings among those. Expressed in terms of the SU(2) and U(1) gauge fields they read

$$\mathcal{L}_{W,Z,A}^{\text{kinetic}} = -\frac{1}{4} W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \;. \tag{1.19}$$

Concerning the role of fermions in the electroweak interactions, it is known that charged currents only involve left-handed particles and right-handed anti-particles,

$$\psi_L = \frac{1}{2}(1 - \gamma^5)\psi, \quad \psi_R = \frac{1}{2}(1 + \gamma^5)\psi, \tag{1.20}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. In the free theory the kinetic terms of the fermions do not mix, $\bar{\psi} i\partial\!\!\!/ \psi = \bar{\psi}_L i\partial\!\!\!/ \psi_L + \bar{\psi}_R i\partial\!\!\!/ \psi_R$. Therefore left and right-handed components can belong to different representations of the gauge group. The latter are assigned to the weak isospin $T = 0$ in order to exclude them from the charged currents and the former are assigned to $T = 1/2$. By assigning different orientations $T^3 = \pm 1/2$ to different flavours of quarks and leptons, respectively, agreement with experimental results can be achieved. For instance, in the first fermion generation

$$L_e = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \;, \quad L_{ud} = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \;. \tag{1.21}$$

With these relations and with the help of the known electric charges of the fermions, $Q = T^3 + Y$, all interior electroweak quantum numbers of the fermion components are fixed. They are presented in Tab. 1.1. Collecting these results, the electroweak interactions of fermions are

$$\begin{aligned}
\sum_L \bar{L} i \not{D} L + \sum_f \bar{f}_{\mathrm{R}} i \not{D} f_{\mathrm{R}} = \; & \mathcal{L}^{\mathrm{free}}_{\mathrm{fermion}} + \mathcal{L}^{\mathrm{int}}_{\mathrm{EW}} \; , \\
\mathcal{L}^{\mathrm{int}}_{\mathrm{EW}} = \; & \frac{e}{\sin\theta_w}(W^+_\mu j^\mu_{W^+} + W^-_\mu j^\mu_{W^-} + Z^0_\mu j^\mu_Z) + e A_\mu j^\mu_{\mathrm{EM}} \; , \qquad (1.22)
\end{aligned}$$

$$\begin{aligned}
j^\mu_{W^\pm} &= \frac{1}{\sqrt{2}} \sum_L \bar{L}^{\pm} \gamma^\mu L^{\mp} \; , \\
j^\mu_Z &= \frac{1}{\cos\theta_w} \sum_f \bar{f} \gamma^\mu (T^3 - \sin^2\theta_w Q) f \; , \\
j^\mu_{\mathrm{EM}} &= \sum_f \bar{f} \gamma^\mu Q f \; . \qquad (1.23)
\end{aligned}$$

L^+ and L^- denote the upper and lower components of fermion doublets, respectively. No right-handed neutrinos are involved in the SM.

Fermion masses

Fermion mass terms couple left and right-handed field components. Therefore they cannot be introduced ad hoc as this would violate gauge invariance due to the different gauge quantum numbers involved. Hence an additional entity must enter these terms. This is the Higgs doublet. For the charged leptons, [3]

$$\begin{aligned}
\mathcal{L}^{\mathrm{mass}}_{\mathrm{leptons}} = \; & - \sum_{i,\, j \,\in\, \{e,\mu,\tau\}} \lambda_{ij} \bar{L}_i \Phi_0 \ell_{R_j} + \mathrm{h.c.} \\
\rightarrow \; & - \sum_{\ell \,\in\, \{e,\mu,\tau\}} \frac{1}{\sqrt{2}} \lambda_\ell v \bar{\ell}_L \ell_R + \mathrm{h.c.} \qquad (1.24)
\end{aligned}$$

Each term of the sum is gauge invariant, and couplings λ_{ij} are introduced. They can always be diagonalized in flavour space by means of chiral rotations $L_i \rightarrow U_{Lij} L_j$ and $\ell_{R_i} \rightarrow U_{Rij} \ell_{R_j}$ with unitary matrices U_L and U_R. Only the mass terms are affected by these rotations. Finally, the lepton mass m_ℓ is given by $m_\ell = \lambda_\ell v/\sqrt{2}$.

[3] "h.c." stands for hermitian conjugate.

1st gen	2nd gen	3rd gen	T	T^3	Y	$Q = T^3 + Y$
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\frac{1}{2}$	$+\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{6}$	$+\frac{2}{3}$ $-\frac{1}{3}$
u_R	c_R	t_R	0	0	$+\frac{2}{3}$	$+\frac{2}{3}$
d_R	s_R	b_R	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$
$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	$\frac{1}{2}$	$+\frac{1}{2}$ $-\frac{1}{2}$	$-\frac{1}{2}$	0 -1
e_R	μ_R	τ_R	0	0	-1	-1
	$\begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix}$		$\frac{1}{2}$	$+\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{2}$	1 0

Table 1.1.: Electroweak quantum numbers of fermions. T and T^3 denote the weak isospin and its third component respectively. Y stands for the weak hypercharge. The electric charge is $Q = T^3 + Y$. Three generations of fermions exist and share the same quantum numbers. In addition there is the Higgs doublet.

A similar mass generation is achieved in the quark sector. However, an additional term is needed to describe massive up-type quarks,

$$\begin{aligned} \mathcal{L}^{\text{mass}}_{\text{quarks}} = & - \sum_{i,\,j\,\in\,\{1,2,3\}} \left(\lambda_d^{ij} \bar{L}^i_{ud} \Phi_0 d_{Rj} + \lambda_u^{ij} \bar{L}^i_{ud} \varepsilon \Phi_0^* u_{Rj} + \text{h.c.}\right) \\ \rightarrow & - \sum_{i\,\in\,\{1,2,3\}} \left(\frac{1}{\sqrt{2}} \lambda_d^i v \bar{d}_{Li} d_{Ri} + \frac{1}{\sqrt{2}} \lambda_u^i v \bar{u}_{Li} u_{Ri} + \text{h.c.}\right) . \end{aligned} \tag{1.25}$$

where $\varepsilon^{ij} = -\varepsilon^{ji}$ and $\varepsilon^{12} = 1$. Again, a diagonalization in flavour space can be realized by chiral rotations, $d_{Li} \rightarrow D_{Lij} d_{Lj}$, $d_{Ri} \rightarrow D_{Rij} d_{Rj}$, $u_{Li} \rightarrow U_{Lij} u_{Lj}$ and $u_{Ri} \rightarrow U_{Rij} u_{Rj}$. The resulting quark fields can thus be interpreted as fermions of distinct masses, but the necessary rotations do not leave every other term of the overall Lagrangian invariant.

The charged W^+ current becomes

$$\begin{aligned} j^{\mu}_{W^+,\text{ quarks}} &= \frac{1}{\sqrt{2}}\bar{u}_{Li}\gamma^{\mu}d_{Li} \\ &\rightarrow \frac{1}{\sqrt{2}}\bar{u}_{Li}\gamma^{\mu}U_{Lij}D^{\dagger}_{Ljk}d_{Lk}\ . \end{aligned} \tag{1.26}$$

The charged W^- current transforms accordingly. The unitary matrix $V_{\text{CKM}} \coloneqq U_L D_L^{\dagger}$ allows for transitions between different generations of quarks [24, 25]. Its nine parameters can be adjusted by means of phase transformations of the quark fields. While there are six arbitrary quark phases, only five CKM parameters can be absorbed in this way since only phase differences enter V_{CKM}. Thus the CKM mixing matrix contains four parameters.

Higgs interactions

So far only the ground state $\Phi = \Phi_0$ has been considered. After the SU(2) gauge fixing is performed, excitations of the Higgs field can be parametrised as follows,

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}\ , \tag{1.27}$$

with $h(x)$ being a real scalar field. It enters the model through the kinetic term according to Eq. 1.18 and through its potential given by Eq. 1.14,

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2}(\partial_{\mu}h)(\partial^{\mu}h) - \frac{1}{2}m_h^2h^2 - \sqrt{\frac{\lambda}{2}}m_h h^3 - \frac{1}{4}\lambda h^4\ . \tag{1.28}$$

Excitations of the Higgs field are scalar particles with a mass $m_h = \sqrt{2\lambda}v$. According to Eq. 1.27, its couplings to other particles follow from the substitution $v \rightarrow v + h$ in the parts of the SM derived above. This means that the Higgs boson couples to all massive particles and only to those. The coupling strength is proportional to the corresponding particle mass. The coupling to the weak gauge bosons is given by

$$\mathcal{L}^{\text{mass}}_{W,Z} + \mathcal{L}^{\text{Higgs}}_{W,Z} = \left(m_W^2W_{\mu}^+W^{-\mu} + \frac{1}{2}m_Z^2Z_{\mu}^0Z^{0\mu}\right)\left(1 + \frac{h}{v}\right)^2\ , \tag{1.29}$$

while the coupling to the massive fermions is given by

$$\mathcal{L}_{\text{fermions}}^{\text{mass}} + \mathcal{L}_{\text{fermions}}^{\text{Higgs}} = -\sum_f m_f \bar{f} f \left(1 + \frac{h}{v}\right) \ . \tag{1.30}$$

Discussion

The SM provides an extensive description of nature. Collecting its different parts, it can be summarized as follows, [4]

$$\begin{aligned}\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{fermions}}^{\text{free}} + \mathcal{L}_{\text{strong}}^{\text{int}} + \mathcal{L}_{\text{gluons}}^{\text{kinetic}} + \mathcal{L}_{\text{EW}}^{\text{int}} + \mathcal{L}_{W,Z,A}^{\text{kinetic}} \\ + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{W,Z}^{\text{mass}} + \mathcal{L}_{W,Z}^{\text{Higgs}} + \mathcal{L}_{\text{fermions}}^{\text{mass}} + \mathcal{L}_{\text{fermions}}^{\text{Higgs}} \ .\end{aligned} \tag{1.31}$$

It accurately describes all known forces of nature except for gravity. In terms of phenomenology a special role is played by QCD. This is due to the behaviour of its coupling strength g_S or $\alpha_S = g_S^2/4\pi$, which needs to be renormalized in order to be able to describe strong interactions at an accuracy which goes beyond the leading order of perturbation theory. This procedure introduces a dependency of the scale at which an interaction occurs. The interpretation of this finding is given by the picture of scattering processes which do not only involve the incoming and outgoing particles but also a multitude of additional surrounding particles, called "vacuum polarisation". The running of the coupling strength with respect to the energy scale Q of a scattering process is determined by the interactions among all of these particles. In the case of QCD it turns out that the coupling strength decreases as the scale Q increases. At the next-to-leading order in perturbation theory one obtains

$$\alpha_S(Q) = \frac{2\pi}{11 - \frac{2}{3} n_f} \frac{1}{\ln \frac{Q}{\Lambda_{\text{QCD}}}} \ . \tag{1.32}$$

where n_f is the number of massless quark flavours [26, 27]. Λ_{QCD} is the scale at which the coupling becomes strong, $\alpha_S(Q) \to 1$. This scale must be determined experimentally and measurements of α_S at any scale can serve this purpose. One obtains $\Lambda_{\text{QCD}} \approx 200\,\text{MeV}$. If one extrapolates Eq. 1.32 to the non-perturbative scale $Q \approx \Lambda_{\text{QCD}}$, the confinement of quarks and gluons in hadrons is explained. The bounding force acting among them becomes larger as their distance grows. As a result, they are always part of bound states.

[4] Once the model is being quantized, additional, gauge-fixing terms are needed. However, this issue goes beyond the scope of this section.

Comprehensive tests of the SM were performed for decades. All of its particles have been discovered by now and nearly all of their properties are in accordance with it. An exception is given by the discovery of neutrino masses. There are more shortcomings of the SM, as briefly mentioned in the introduction. Therefore further tests of the SM are motivated. Many models of physics beyond the SM have been constructed and now, apparently, only experimental results can guide a clear path towards an understanding of nature which goes beyond the SM. There are indications that such phenomena will matter at the TeV scale, which is accessible at the LHC. However, no sign of new phenomena has been found in LHC collisions to date. Careful studies of this data must go on, be it in the context of searches for specific new phenomena, or in the context of tests of selected parts of Eq. 1.31, which is the aim of this thesis.

1.2. Hadronic collision phenomenology

1.2.1. The factorization theorem

Hadronic collisions at high energies are described by the parton model. Scatterings occur among the parton constituents of hadrons, namely quarks and gluons, which are asymptotically free (Eq. 1.32). The momentum fraction x of partons with respect to their parent hadron, $p_{\text{Parton}} = xP_{\text{Hadron}}$, well-defined in a frame of large hadron momentum, is distributed according to universal probability densities $f(x)$, called parton distribution functions (PDFs). They allow one to factorize the soft interactions of a hadron on the one hand and the hard interactions of partons on the other hand. Accordingly, hadronic cross sections are computed as convolutions of PDFs with partonic cross sections. The asymptotic freedom of the partons is reflected by the invariance of the PDFs with respect to the energy scale of the hard scattering.

Indeed, the parton model successfully describes hadronic collisions at high energies. However, a closer inspection reveals problems, both experimentally and theoretically. Experimentally, the scale invariance of PDFs is violated. They depend logarithmically on the energy scale of the hard interaction. Theoretically, the problem is that in perturbation theory beyond the leading order divergences arise in the computations of hard scattering cross sections, which are caused by collinear radiation in the initial state and do not cancel in the context of perturbation theory. As it turns out, there is a common solution to these problems which leads to the QCD improved parton model.

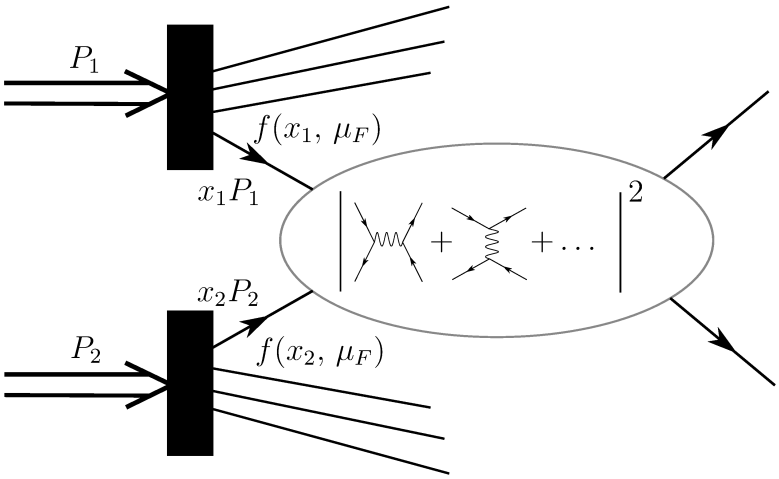


Figure 1.3.: Illustration of the factorization theorem. In high energy collisions of hadrons, scatterings occur among their parton constituents. The longitudinal momentum fractions of the partons are distributed according to the PDFs $f(x, \mu_F)$. The hard scattering cross section is factorized off the hadronic part and can be calculated by means of perturbation theory.

Concerning the violation of the energy scale invariance, it is important to notice that partons are only asymptotically free at high scales Q. Residual strong interactions alter the PDFs depending on Q. If for a given value of x a higher scale $Q+\Delta Q$ is probed, partons of higher values of x can contribute to the PDF at the given lower value of x due to the radiation of partons. This radiation can be described by QCD, which leads to evolution equations of PDFs, called DGLAP equations [28–30],

$$\frac{\mathrm{d}}{\mathrm{d}\ln Q^2}\begin{pmatrix} f_i(x,Q) \\ f_g(x,Q) \end{pmatrix} = \alpha_S(Q^2) \sum_{j\,\in\,\{q,\bar{q}\}} \int_x^1 \frac{\mathrm{d}z}{z} \begin{pmatrix} P_{q_iq_j}(\frac{x}{z}) & P_{q_ig}(\frac{x}{z}) \\ P_{gq_j}(\frac{x}{z}) & P_{gg}(\frac{x}{z}) \end{pmatrix} \begin{pmatrix} f_j(z,Q) \\ f_g(z,Q) \end{pmatrix} . \tag{1.33}$$

They convolute PDFs at momentum fractions $z > x$ with splitting functions P at splitting fractions x/z, which results in changes of PDFs at the momentum fraction $z \cdot x/z = x$. The functions P are computed by means of perturbative QCD. Once the PDFs are determined at a certain scale experimentally, the DGLAP equations allow one to determine them at any other scale as long as perturbation theory holds ($Q > \Lambda_{\mathrm{QCD}}$).

The theoretical problem of collinear divergences which appear in hard scattering cross section calculations is caused by the fact that PDFs refer to the dynamics of hadrons. Due to confinement, they cannot be described by perturbative QCD. However, differences of PDFs at different high scales are not affected by confinement. Accordingly, PDFs

can be renormalized. This procedure introduces a PDF renormalization scale, called the factorization scale μ_F. The renormalization group equations of PDFs are the DGLAP equations 1.33. The renormalization scheme and scale chosen with respect to the PDFs must be taken into account in the calculation of the hard scattering cross section in order to avoid a double-counting of parton radiation.

In summary, hadronic collisions can be factorized into soft components described by PDFs $f_i(x, Q = \mu_F)$ and hard scattering cross sections $\hat{\sigma}(\hat{s}, \mu_R, \mu_F)$ described by perturbative QCD. This relation is illustrated in Fig. 1.3. For instance, total proton-proton cross sections are given by

$$\begin{aligned} \sigma(s) &= \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i(x_1, \mu_F) f_j(x_2, \mu_F) \hat{\sigma}_{ij}(\hat{s}, \mu_R, \mu_F) \ , \\ \hat{s} &= (x_1 P_{\mathrm{Proton\,1}} + x_2 P_{\mathrm{Proton\,2}})^2 \\ &= x_1 x_2 s \ , \end{aligned} \tag{1.34}$$

up to small non-factorizable contributions. A sum over initial state partons i, j is performed. The hadronic and partonic centre-of-mass energies squared are denoted by s and $\hat{s}$, respectively. The renormalization scale is denoted by μ_R. Usually the scales μ_R and μ_F are chosen to be equal to a typical scale of the scattering process in question, for instance the mass of a heavy particle, which can help to avoid large logarithms in perturbative expansions. While the hadronic cross section is ultimately independent of these scales, a finite dependency remains at any fixed order in perturbation theory. Scale variations are formally of a higher order and are used to roughly estimate the accuracy of predictions.

The factorization theorem allows one to predict hadronic collision rates. Essentially it states that the physics at high scales, e.g. $Q \approx 100\,\mathrm{GeV}$, described by $\hat{\sigma}$, do not depend on the physics at the lower nuclear scale $Q \approx 1\,\mathrm{GeV}$ described by the PDFs. An analogue of this theorem is the observation that the physics of a proton does not depend on the physics of a possible atomic shell surrounding it, as the atomic scale is only $Q \approx 10\,\mathrm{eV}$.

This theorem can be used in order to interpret measurements in terms of PDFs. For instance, PDFs at a certain scale can be parametrized and the parameters are extracted from fits to experimental data. Then the resulting PDFs can be used to make predictions for other experiments. Figure 1.4 shows an example of a PDF set determined by the MSTW PDF fitting group [31], where two PDFs are compared with each other at two different energy scales. As higher scales are probed, smaller values of x become

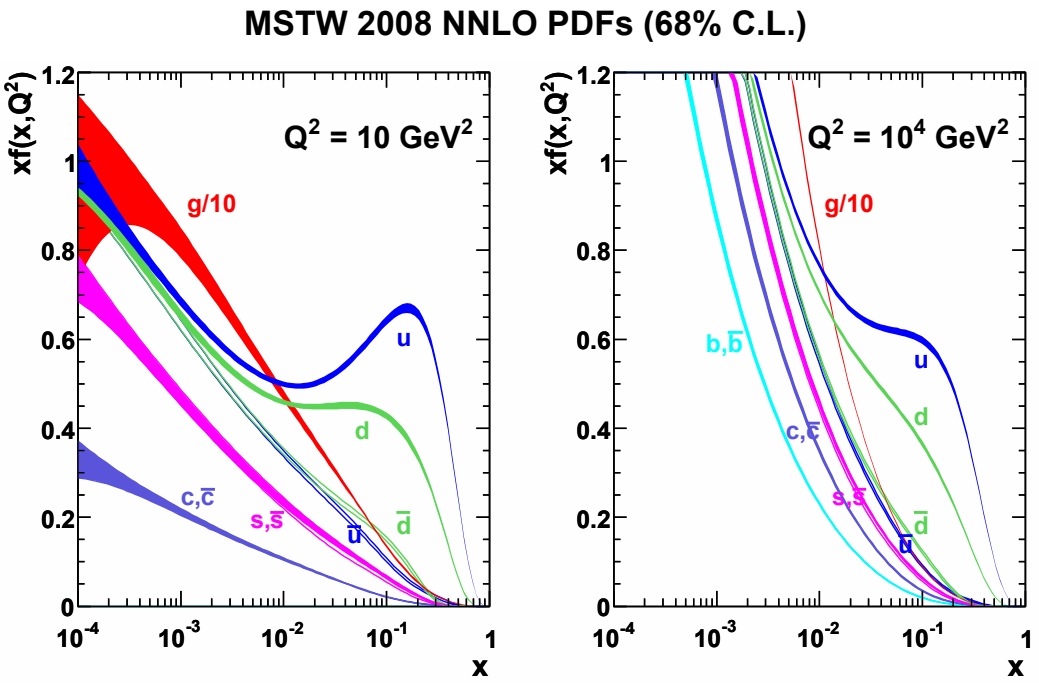


Figure 1.4.: MSTW2008 NNLO proton PDF set. The product $xf(x,Q)$ is shown for the different kinds of partons. The valence quarks u and d are abundant at high momentum fractions x. The gluon PDF is particularly prominent. See-quarks accumulate at low x values. At higher energy scales Q^2, more parton radiation is resolved. In consequence, all PDFs grow at low x values [32].

more prominent since more parton radiation is resolved. In general, PDF fits and the employment of PDFs involve the DGLAP equations 1.33 which are evaluated at a certain order in α_S. This order should not be lower than the order involved in the calculation of $\hat{\sigma}$ in order to reach the desired precision.

1.2.2. Jets of strongly interacting particles

While in the previous section the parton model of strong interactions was used to describe the initial state of hadronic collisions leading to hard scattering processes, this section is dedicated to the role of strong interactions in the final state. The description of hard scatterings in the context of perturbation theory is well suited for the involved high energy scales. However, once partons of final states become separated from each other and from the hadron remnants, the strong interaction actually becomes strong and colour singlet hadrons are formed. At this point perturbation theory is not useful anymore. Still,

final state hadrons reflect properties of final state partons. At high energies, collimated bunches of hadrons are observed and can be associated with quarks and gluons described by the parton model.

In order to understand the formation of these jets, algorithms are used which combine single particles into jets. They can be employed on the level of hadrons or energy deposits in particle detectors in order to obtain experimental jet cross sections, as well as on the parton level in order to obtain theoretical jet cross sections to compare with. As jets are a major aspect throughout this thesis, a short description of jet algorithms is presented in the following.

Two desired properties of jet algorithms are derived from the requirement of well-defined jet cross sections at the parton level. Beyond the leading order of perturbation theory, loop corrections and bremsstrahlung enter the cross section calculations. Both of them involve singularities. These are related to the emission of soft and collinear particles, which cannot be distinguished experimentally. For sufficiently inclusive observables, the singularities cancel as soon as all contributions are summed up [33, 34].

This condition of insensitivity to soft and collinear radiation is met by certain jet clustering algorithms. Here a distance measure $d_{i,j}$ between pairs of outgoing particles (i, j) is defined together with a measure of the hardness of each particle i itself, denoted by d_i. In each of the sequential clustering steps, the minimum of all measures is identified. If it is a $d_{i,j}$ measure, the particles i and j are combined and treated like a single particle in the next clustering step. If it is a d_i measure, the particle i is declared to be a jet and removed from the list of clustering input. This procedure is repeated until all particles became part of a jet. An example of an infrared and collinear safe jet clustering procedure is the anti-k_T clustering algorithm [35], which is most popular at the LHC experiments. It is defined by the following equations,

$$d_{i,j} \coloneqq \min\left(\frac{1}{{k_\mathrm{T}}_i^2}, \frac{1}{{k_\mathrm{T}}_j^2}\right) \cdot \frac{\Delta R_{i,j}^2}{R^2}\,, \tag{1.35}$$

$$d_i \coloneqq \frac{1}{{k_\mathrm{T}}_i^2}\,. \tag{1.36}$$

Here k_T denotes the transverse momentum and $\Delta R_{i,j} = \sqrt{(\Delta\eta_{i,j})^2 + (\Delta\varphi_{i,j})^2}$ denotes a geometrical distance between the particles i and j with $\Delta\varphi_{i,j}$ and $\Delta\eta_{i,j}$ being the differences of azimuthal angles and pseudorapidities, respectively. Pseudorapidities are proportional to polar angles and will be discussed in Sec. 2.2. The parameter R denotes

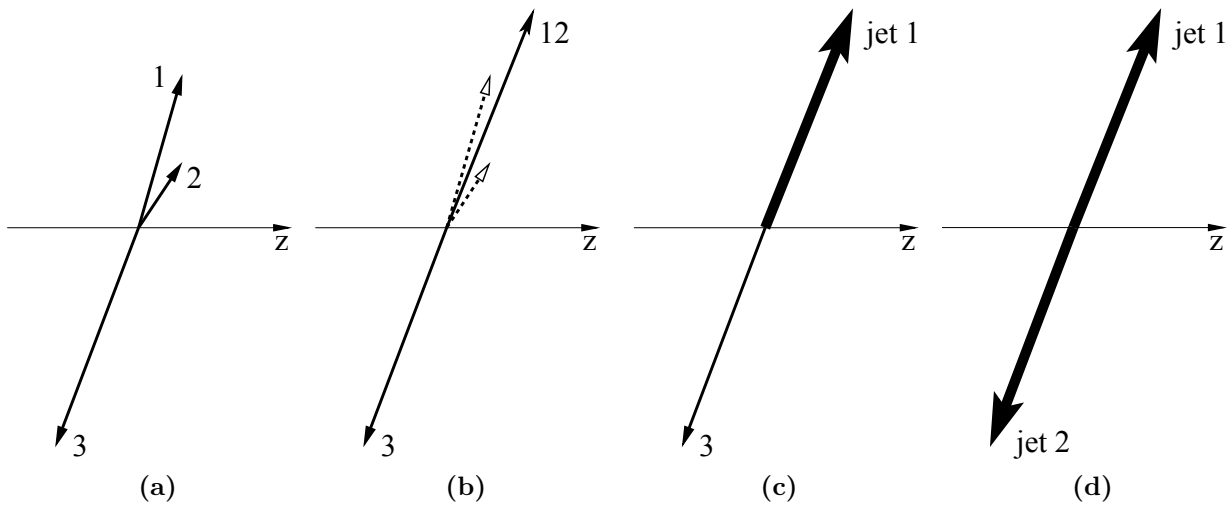


Figure 1.5.: Jet clustering with the anti-k_T algorithm, as defined by Eq. 1.36. Final state particles are the input to the clustering procedure (a). In the first step the particles 1 and 2 are merged because k_{T1} is large and $\Delta R_{1,2}$ is small (b). In the second step the merged object 12 is not merged again but declared to be a jet because $\Delta R_{12,3}$ is large (c). Finally, the remaining particle 3 is declared to be a jet (d). Thus the three particle configuration is merged into a dijet final state.

a fixed distance, usually chosen to be about 0.4 . This algorithm typically considers the hardest particle first, combining it with softer particles in its vicinity $\Delta R_{i,j}/R < 1$. Hence it basically reverses the emission process and allows one to relate various softer final state objects to fewer primary objects of the hard scattering. An example of such a clustering procedure is shown in Fig. 1.5. The anti-k_T algorithm is insensitive the presence of soft and collinear particles. They are always clustered with other particles due to their small transverse momentum, $k_T \approx 0$, or due to their small distance to another particle, $\Delta R \approx 0$.

1.2.3. Event rates at hadron colliders

At hadron colliders, in principle all of the known fundamental interactions except for gravity can be studied due to the interactions of quarks and gluons at high energies. Electroweak interactions are possible due to the presence of quarks. Higgs interactions are possible mainly through couplings to intermediate heavy particles. Strong interactions are dominant though. Among these, soft QCD processes which lead to final state objects of low transverse momenta p_T have the highest interaction rates, followed by the

production of jets in strong interactions. Only relatively small cross sections are obtained for electroweak and Higgs processes, as shown in Fig. 1.6. In consequence, most of the experimental analyses need to deal with backgrounds caused by strong interactions. In particular strong interactions frequently result in low scattering angles as well as low transverse momenta. Accordingly, analyses of electroweak processes are often restricted to a subset of the overall phase space which is smaller than the available region of experimental acceptance and characterized by central, high p_{T} final state objects.

1.3. Top-quark phenomenology

The top-quark was the sixth and last quark flavour to be discovered, namely by the CDF and D0 collaborations at the Fermilab Tevatron $p\bar{p}$ collider in 1995 [37, 38]. Its existence was already anticipated much earlier, once the b-quark and the τ-lepton were discovered. Aside from intuition, there is a strict theoretical reason why also the third generation of fermions is supposed to come with an up-type quark. It is related to anomalies in QFTs, where a symmetry which is present at the level of classical fields does not hold once the theory is quantized. For instance, there is the violation of scale invariance in QCD as discussed in Sec. 1.2.1, even in the case of massless quarks where no scale is involved in the classical field theory. In the case of classically conserved currents coupling to massless vector bosons, anomalies give rise to violations of unitarity. Thus a QFT must be free of such anomalies in order to be consistent. The standard model is only free of these gauge anomalies if the generations of quarks and leptons are complete (Tab. 1.1). Furthermore, the requirement of the absence of gauge anomalies relates the fermion quantum numbers to each other.

Experimentally, thorough studies of electroweak precision observables related to the b-quark gave hints to the quantum numbers of the top-quark. In particular, measurements of the b-quark decay width of the Z boson and the forward-backward asymmetry $A_{\mathrm{FB}}(b)$ at e^+e^- colliders, interpreted within the standard model, lead to $T_L^3(b) = -1/2$ and $T_R^3(b) = 0$ for the left and right-handed components, respectively, which suggests the existence of a weak isospin partner of the b-quark [39]. Furthermore, the observation of $B_d^0\bar{B}_d^0$ mixing at the $\Upsilon(4s)$ resonance revealed a circumvention of the GIM mechanism which could be explained by the action of a sixth quark with a high mass [40, 41]. Also the top-quark mass can be predicted in the context of fits to electroweak precision observables, in particular to measurements at the Z boson pole and to the W boson mass and width, since top-quark couplings to the massive vector bosons alter the boson

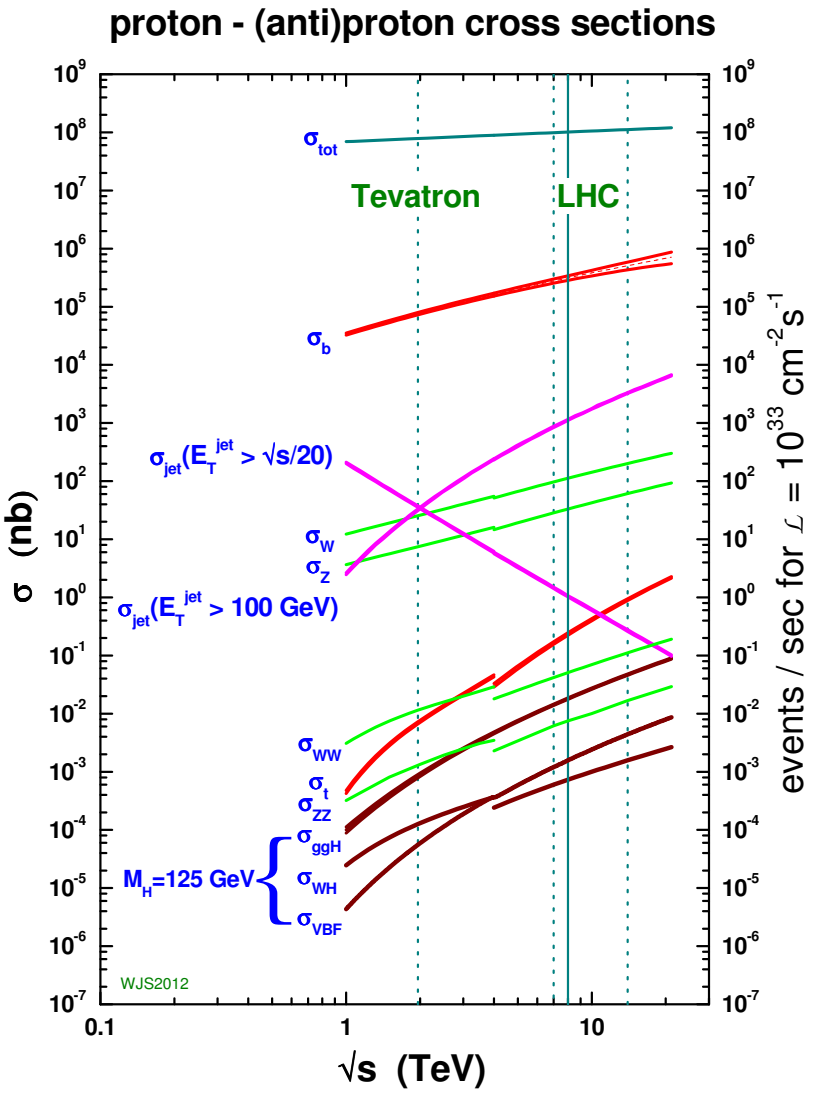


Figure 1.6.: Cross sections at hadron colliders. Total cross sections of $p\bar{p}$ (Tevatron) and pp (LHC) collisions are shown on the left and right-hand side, respectively. Processes of strong interactions are dominant. Electroweak processes need to be distinguished thoroughly from them [36].

masses. A corresponding fit results in $m_t = 180 \pm 10$ GeV [42], which agrees with the measurement of $m_t = 173.34 \pm 0.76$ GeV [7].

Its high mass makes the top-quark a particularly interesting object for two reasons. First of all, it decays before it can hadronize since a large decay phase space is available. Therefore it offers a unique possibility to study an almost bare quark. Secondly, the forces responsible for electroweak symmetry breaking can be expected to couple strongly

to the massive top-quark. In the following, an overview of the phenomenology of this special particle is presented where emphasis is put on the LHC environment [43].

1.3.1. Top-quark decay

Within the SM, the top-quark decays almost exclusively via charged currents into a W boson and a down-type quark, $t \to W q_d$. Decays via neutral currents involve loops of intermediate particles and are negligible. Similar relations hold for the anti-top-quark. Due to its high mass, the top-quark decays particularly rapidly. In the leading order of perturbation theory, its total decay width is

$$\Gamma_t^{\mathrm{LO}}(t \to W q_d) = \frac{e^2}{64\pi \sin^2 \vartheta_W}(|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2)\frac{m_t^3}{m_W^2}\left(1 - \frac{m_W^2}{m_t^2}\right)^2\left(1 + 2\frac{m_W^2}{m_t^2}\right) , \tag{1.37}$$

where the masses of the down-type quarks are neglected and the W boson is on the mass shell. The sum of CKM-matrix elements $|V_{tq_d}|^2$ in brackets equals one within the SM, where the 3×3 CKM matrix is unitary. More precisely, at next-to-leading order in QCD the total decay width of the top-quark amounts to $\Gamma_t = 1.35\,\mathrm{GeV}$, which corresponds to a lifetime of about $5 \cdot 10^{-25}\,\mathrm{s}$ [44]. As this lifetime is shorter than the typical time scale of hadronic interactions, $1/\Lambda_{\mathrm{QCD}} \approx 1/200\,\mathrm{MeV} = 3 \cdot 10^{-24}\,\mathrm{s}$, top-quarks do not form bound states.

The determination of the CKM matrix elements $|V_{tq_d}|$ is subject to ongoing research as discussed below. Within the SM, the unitarity of the 3×3 CKM matrix allows one to infer their values from other CKM matrix elements determined in dedicated experiments. This approach results in $|V_{tb}| = 0.99915^{+0.00005}_{-0.00004}$ [45]. Hence the top-quark is expected to decay predominantly into a b-quark and a W boson. Subsequently, the W boson decays into quarks or leptons. The branching ratio for each of the leptonic decays is $\mathcal{B}(t \to b\ell^+\nu_\ell) = 0.108$, $\ell \in \{e, \mu, \tau\}$ [43]. In turn, 68% of all top-quark decays are purely hadronic.

The angular distributions of top-quark decay products are determined by the $V-A$ structure of the charged currents and by angular momentum conservation. The high mass of the top-quark converts into a high kinetic energy of the b-quark which results from its decay. Therefore the b-quark mass can be neglected and the $V-A$ coupling structure leads to b-quarks of negative helicity, which in turn constrains the allowed

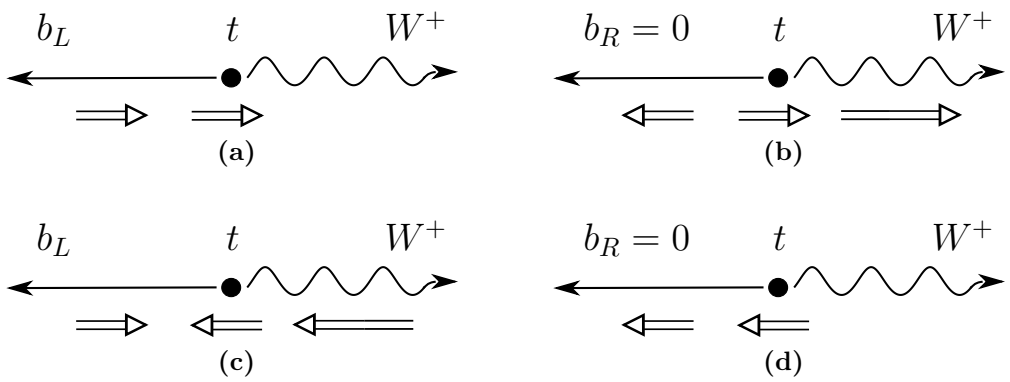


Figure 1.7.: Top-quark decay in the top-quark rest frame at the leading order of perturbation theory. The spin quantization axis is chosen to be aligned with the momenta of the resulting b-quark and W boson. In (a) to (d) all decays in accordance with angular momentum conservation are shown. Due to the $V-A$ structure of charged currents, only b-quarks with negative helicity are allowed in the limit $m_b \to 0$, as shown in (a) and (c). In consequence, only longitudinally and negatively polarized W bosons occur.

helicities of the W boson to be either longitudinal or negative. Figure 1.7 illustrates all decay scenarios in leading order perturbation theory. Here the helicity fractions of the W boson are $F_0 = 0.7$, $F_- = 0.3$ and $F_+ = 0$. Beyond the leading order these relations hold up to the per cent level. Thus the SM makes a distinct prediction of the topology of top-quark decays, which can be tested for it is not smeared by hadronization effects. In particular, the abundance of longitudinally polarized W bosons is of interest, as these states are a consequence of electroweak symmetry breaking [43].

1.3.2. Top-quark production in hadronic collisions

Due to its high mass, the top-quark could not be produced by electron-positron colliders to date. Only hadron colliders could provide collisions of sufficiently high energies. In addition, high luminosities are required in order to produce a number of top-quarks which is large enough to allow for the detection of a significant signal. The first collider to reach these goals was the Tevatron $p\bar{p}$ collider at Fermilab, followed by the pp collider LHC at CERN [46, 47]. There are two relevant production modes, namely the pair production via strong interactions and the production of single top-quarks via electroweak interactions.

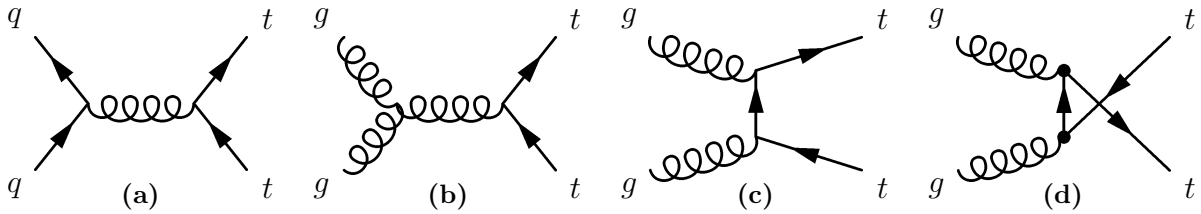


Figure 1.8.: Leading order Feynman diagrams of top-quark pair production. Quark-anti-quark annihilation, (a), and several channels of gluon fusion, (b) to (d), contribute.

Top-quark pair production

The pair production of a top and an anti-top-quark via strong interactions is the dominant production mode of top-quarks at hadron colliders. Concerning the hard scattering process, quark-anti-quark annihilation and gluon fusion contribute (Fig. 1.8) while quark-gluon scattering enters beyond the leading order of perturbation theory. As anti-quarks seldomly contribute at the pp collider LHC, gluon fusion is the dominant production mode in this case. It contributes about 90% of the total hadronic $t\bar{t}$ cross section, depending on μ_F [43]. Due to this large gluon contribution, the lack of an anti-proton beam at the LHC is basically no disadvantage in the case of $t\bar{t}$ production. The total hadronic $t\bar{t}$ production cross section has been calculated up to the next-to-next-to-leading order (NNLO) in QCD. In addition, logarithms which become large in the region of low invariant $t\bar{t}$ mass, $M_{t\bar{t}} \approx \sqrt{\hat{s}}$, have been summed, giving next-to-next-to-leading logarithmic (NNLL) corrections [48–54]. In order to determine the theoretical cross section uncertainty, variations of the scales μ_R and μ_F are examined and uncertainties of PDFs and the strong coupling α_S are propagated. Using the PDF sets MSTW2008 NNLO [31, 55], CT10 NNLO [56, 57] and NNPDF2.3 5f FFN [58] in accordance with the PDF4LHC prescription [59] to be discussed in Sec. 1.3.4, the total $t\bar{t}$ cross section at the LHC for a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ is $\sigma_{pp\to t\bar{t}+X} = 253^{+13}_{-15}\,\text{pb}$, given a top-quark mass of $m_t = 172.5\,\text{GeV}$. Concerning the pp collision data sets recorded by the ATLAS and CMS experiments in 2012, which correspond to an integrated luminosity of about $L = 20\,\text{fb}^{-1}$ in each case, this cross sections results in the production of about 5 million top-quark pairs.

The experimental signature of $t\bar{t}$ production depends on the decay mode of the W bosons produced in the decays of the two top-quarks as shown in Fig. 1.9. Three channels are distinguished, which are the di-lepton, the single lepton and the all hadronic channel, where two, one or none of the W bosons decay into leptons, respectively. While

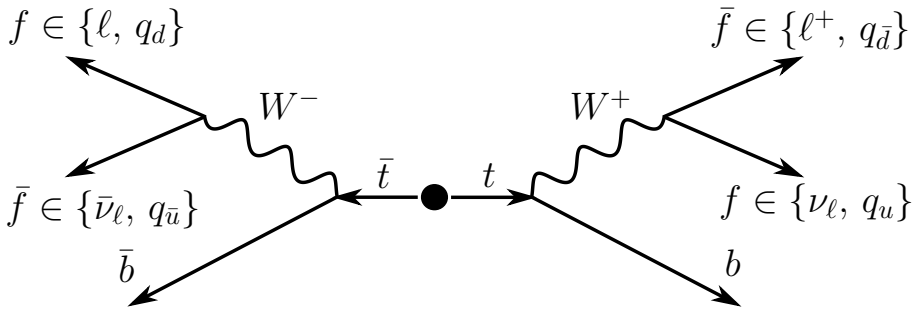


Figure 1.9.: Top-quark pair production event topology. The experimental signature of $t\bar{t}$ events is determined by the decay modes of the two W bosons resulting from the top and anti-top-quark decays. The branching fractions of the all-hadronic, single lepton and di-lepton channels are 46%, 44% and 11%, respectively [44].

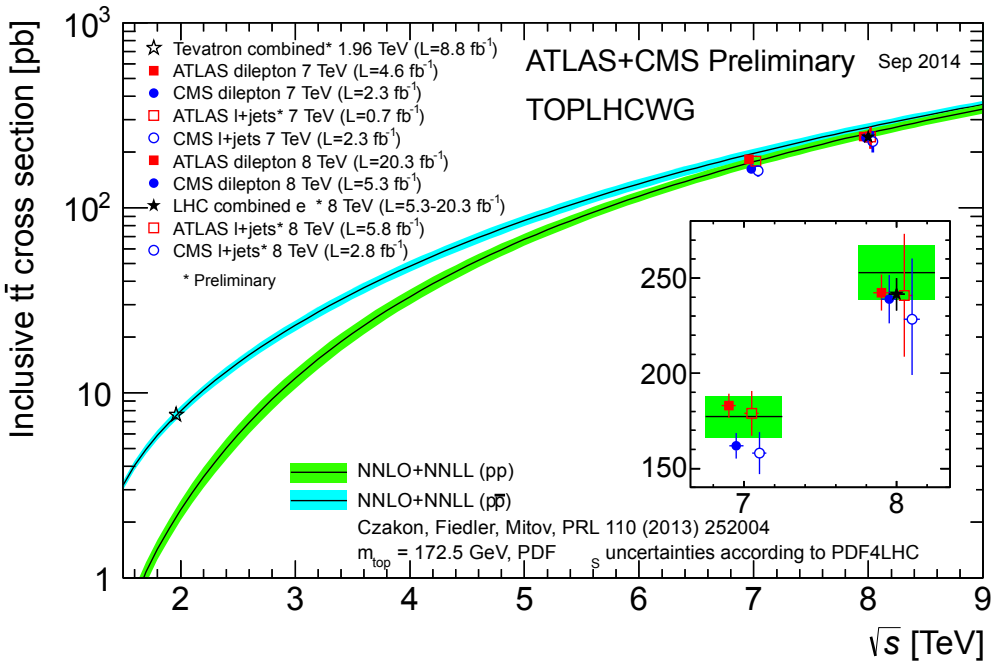


Figure 1.10.: Top-quark pair production cross sections. Various measurements have been performed by the CDF and D0 collaborations at the Tevatron and by the ATLAS and CMS collaborations at the LHC. They agree with the SM predictions [60].

the all hadronic channel competes with large backgrounds caused by jets produced in strong interactions, the leptonic channels are relatively clean and provide electrons and muons which are convenient trigger objects. For these reasons, studies of the top-quark mostly focus on the two leptonic channels. Aside from the charged leptons, $t\bar{t}$ collision events are characterized by multiple jets, up to two of which can be associated with b-quarks. Given the long lifetime of the B hadrons into which the b-quarks hadronize and the sizeable boost of the b-quarks, caused by the large mass of the decaying top-quarks, secondary decays of B hadrons can be identified. Another sign of leptonic top-quark decays is given by undetected or missing transverse momentum due to neutrinos. In general, large transverse momenta are involved.

Meanwhile experimental techniques have been established and improved up to a level of accuracy which surpasses the precision of the theoretical predictions. Figure 1.10 presents various measurements performed at the Tevatron and at the LHC. They agree with the theoretical predictions. This level of understanding allows one to go beyond total $t\bar{t}$ cross section measurements. Aside from measurements of top-quark properties in $t\bar{t}$ events, it also allows one to measure processes which involve $t\bar{t}$ production as a background, including electroweak top-quark production.

Single top-quark production

Top-quarks can also be produced singly. As the top flavour changes in these reactions, they must involve charged currents. In consequence, the phenomenology of single top-quark production differs distinctly from the case of top-quark pair production. In particular, CKM matrix elements are involved not only in the decay of the top-quark but also in its production. Furthermore, high degrees of top-quark polarisation are obtained, which translate into distinct topologies of final states.

Three production modes are distinguished, namely the s-channel, the t-channel and the associated production of a top-quark and a W boson. Their leading order Feynman diagrams are shown in Fig. 1.11. The s and t-channel are distinguished by the exchanged W boson being either timelike or spacelike, respectively. They do not interfere with each other up to NNLO in QCD because of their different colour structure. The t-channel can be described either in the four or five flavour scheme. In the latter case, a b-quark PDF is employed by means of the DGLAP evolution equations 1.33. The Wt mode is another way to produce single top-quarks. However, the initial and final states of its leading order diagrams, Fig. 1.11d and 1.11e, are similar to the top-quark pair production

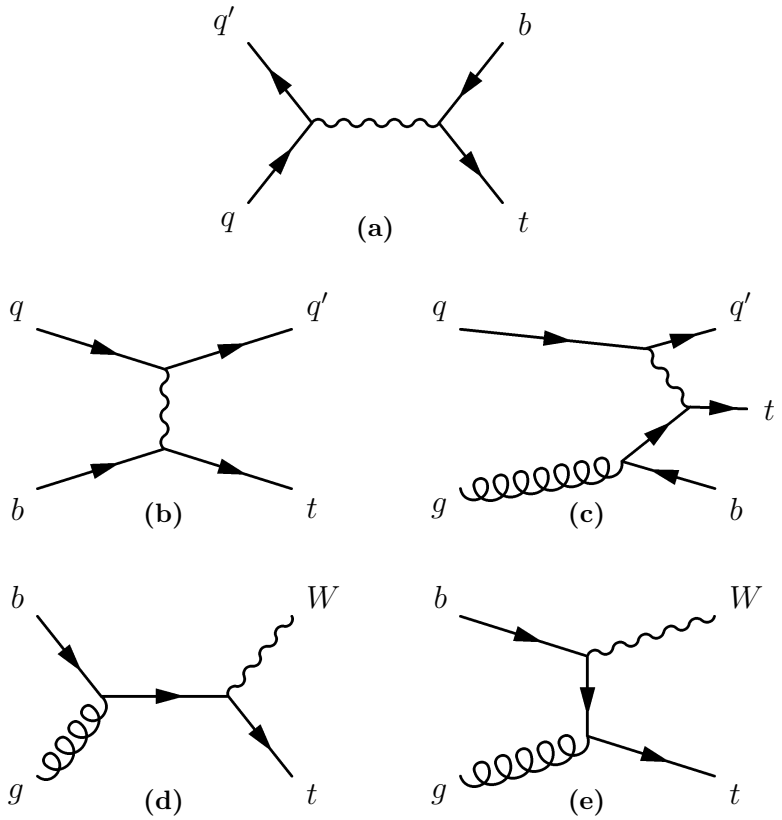


Figure 1.11.: Leading order Feynman diagrams of single top-quark production. The s-channel is characterized by the exchange of a timelike W boson (a), while it is spacelike in case of the t-channel. The latter process can either be described in the five (b) or four flavour scheme (c). The associated production of a top-quark and a W boson also involves the Wtb vertex (d), (e). In general, s and d-quarks can be involved in the top-quark coupling as well.

process with scatterings $gg \to t\bar{t}$ and $\bar{t} \to W^-\bar{b}$. Already at NLO in QCD these processes interfere. Still, collision events can be identified with Wt single top-quark production if they are subject to appropriate kinematic selection cuts [61].

Being electroweak processes, the cross sections of single top-quark production are smaller than those of top-quark pair production. However, due to the lower production threshold the difference is relatively small. At the LHC with $\sqrt{s} = 8\,\text{TeV}$ the t-channel cross section is $\sigma^t_{pp\to tq+X} = 85^{+4}_{-3}\,\text{pb}$, followed by Wt production with $\sigma_{pp\to Wt+X} = 22 \pm 2\,\text{pb}$. The s-channel cross section is particularly small at the pp collider LHC due to the demand of initial anti-quarks, $\sigma^s_{pp\to tb+X} = 5.2 \pm 0.2\,\text{pb}$. The s and t-channel cross sections are determined at NLO in QCD [62–70]. The uncertainties include variations of the renormalization and factorization scales as well as PDF uncertainties including variations of the strong coupling. Concerning PDFs, the PDF4LHC prescription is applied [59], which is discussed further in the next section. The

Wt production cross section is based on approximate NNLO calculations [71] using the MSTW2008 NNLO PDF set. Here PDF uncertainties are evaluated using the 90% confidence level PDF variation provided by the MSTW group. In addition, renormalization and factorization uncertainties are taken into account. A top-quark mass of 172.5 GeV is chosen throughout the calculations. These cross sections are in accordance with the recommendations of the LHC Top Working Group, which aims for a harmonisation of choices of reference top-quark cross sections [72]. The s-channel cross section prediction is tested by the measurement presented in this thesis. For this reason a detailed discussion of single top-quark cross section predictions is presented in the next section, while the present section is focused on the phenomenology of both the s-channel and the t-channel single top-quark production, as the latter is one of the major background processes regarding the former.

Measurements of s and t-channel single top-quark production are restricted to leptonic top-quark decays for the same reasons as in the case of top-quark pair production. Furthermore, the presence of two or three jets is required, up to two of which are identified to be induced by a b-quark. While t-channel event selections aim for one b-tag, two b-tags are often required in s-channel selections, in accordance with the generic signature shown in Fig. 1.11a with the decay $t \to Wb$. A-priori, these b-tagging requirements reduce the sensitivity to scenarios beyond the SM with sizeable CKM matrix elements $|V_{td}|$ and $|V_{ts}|$, but they are necessary in order to discriminate the signals from abundant background processes of W and Z boson production in association with jets. In addition, the pair production of top-quarks is another considerable background process, which restricts single top-quark measurements to low jet multiplicities. Even then backgrounds give sizeable contributions, so that usually advanced techniques need to be applied in order to identify a signal. In this way, the first observation of single top-quark production was achieved by the CDF and D0 collaborations at the Tevatron in 2009 [73, 74]. Here the sum of s and t-channel contributions was considered as one signal. Separate observations followed later, including the observation of s-channel single top-quark production by means of a combination of CDF and D0 measurements in 2014 [75].

Indeed it is important to disentangle the s and t-channel, because contributions of physics beyond the SM affect them differently [76]. For instance, flavour changing neutral currents (FCNCs) like ugt or cgt couplings allow for an enhancement of single top-quark cross sections. In the s-channel, FCNCs result in final states like $t\bar{u}$, $t\bar{c}$ instead of $t\bar{b}$. In consequence, s-channel measurements are likely to miss FCNC contributions given the

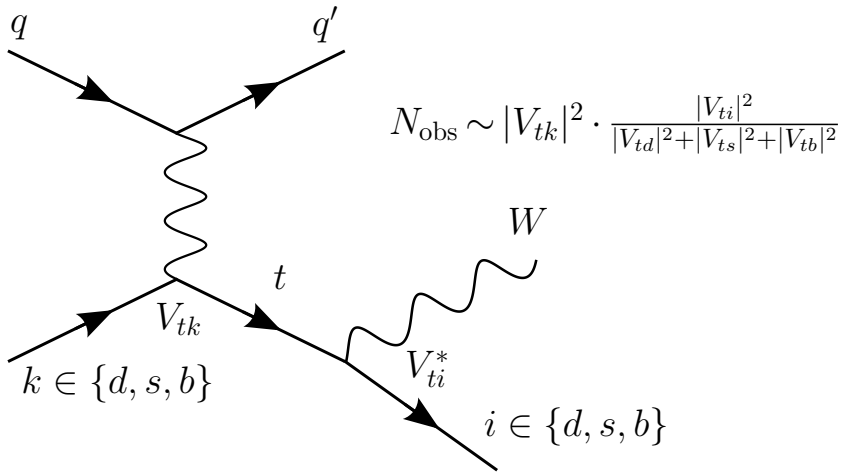


Figure 1.12.: CKM matrix elements in single top-quark production and decay. Here the t-channel is shown as an example. The number of observed events N_{obs} is proportional to CKM matrix elements squared which enter the top-quark production and a ratio of CKM matrix elements related to the top-quark decay.

requirement of two b-tagged jets discussed above. On the other hand, there is sensitivity to FCNCs in t-channel measurements. Here the cross section can be enhanced by FCNCs as well, in particular because of the larger contributions of up and charm-quark PDFs compared to the b-quark PDF involved in the SM process. The same effect matters if new phenomena result in an enhancement of the CKM matrix elements $|V_{td}|$ and $|V_{ts}|$ compared to the SM expectation $|V_{td}|, |V_{ts}| \ll 1$, again enhancing the t-channel rate. The observed s-channel rate, on the other hand, is reduced in this scenario as $|V_{tb}|$ would be less than its SM expectation $|V_{tb}| \approx 1$. Other extensions of the SM which are relevant to single top-quark production predict heavy charged gauge bosons, usually denoted W', which strongly couple to the top and b-quarks. In particular they contribute to the s-channel rate where a resonant W' production is possible. In summary, it is necessary to disentangle the s and t-channel processes in order to gain sensitivity to different scenarios of physics beyond the SM.

1.3.3. CKM matrix elements

A major motivation for single top-quark measurements is given by its sensitivity to the CKM matrix elements $|V_{td}|$, $|V_{ts}|$ and $|V_{tb}|$. In single top-quark processes they enter the production and the decay of the top-quark. In top-quark pair production via strong interactions, on the other hand, only top-quark decays involve these matrix elements. Experimentally, top-quark pair production is best suited for a measurement of the

branching fraction of the top-quark into a W boson and a b-quark. This is because jets induced by b-quarks can be tagged as such efficiently as mentioned above, while this is hardly achievable in case of s and d-quarks (a discussion of b-tagging is given in Sec. 2.5.4). Neglecting the masses of the down-type quarks, this ratio is given by

$$R_b = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} \ . \tag{1.38}$$

It can be measured on the basis of b-tag multiplicities in collision events enriched with top-quark pair production. Its most precise measurements to date are provided by the D0 collaboration yielding $R_b = 0.90 \pm 0.04$ [77] and by the CMS collaboration yielding $R_b = 1.014 \pm 0.032$ [78], which constitutes a slight tension with respect to the former result. In single top-quark measurements R_b enters as well, as usually final states including b-quarks are considered. In addition, a CKM matrix element enters due to the production of top-quarks via weak currents, Fig. 1.12. Experimentally, the CKM matrix elements in question can be related to measured rates of top-quark pair and single top-quark production together with selection efficiencies of the particular production and decay channels. This approach is model-independent except for the assumption of SM coupling structures [79].

In practice, such a measurement has not been performed yet. Instead, measurements of R_b are usually interpreted in terms of a measurement of $|V_{tb}|$ using the relation $|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 = 1$, derived from 3×3 CKM unitarity. However, given this assumption $|V_{tb}|$ is tightly constrained already by means of other measurements of the flavour sector. In other words, the R_b measurements only prove that $|V_{tb}|$ is greater than $|V_{ts}|$ and $|V_{td}|$. Concerning single top-quark measurements, it is usually assumed that $R_b = 1$, or equivalently $|V_{tb}| \gg |V_{td}|, |V_{ts}|$, so that Wtd as well as Wts production vertices do not matter and all top-quarks decay via $t \to Wb$. In turn, single top-quark cross sections are proportional to $|V_{tb}|^2$. Again, the $|V_{tb}|$ measurement is quite model-dependent. A combination of CDF and D0 measurements yields $|V_{tb}| = 1.02^{+0.06}_{-0.05}$ [80]. At the LHC, ATLAS reports a measurement of $|V_{tb}| = 1.02 \pm 0.07$ [81], while CMS reports $|V_{tb}| = 0.998 \pm 0.041$ [82]. Still, a less model dependent determination remains to be performed. All of the aforementioned modes of top-quark production can provide important input to it, including s-channel single top-quark production.

1.3.4. Single top-quark cross section predictions

Numerous theoretical studies of single top-quark production are available with respect to all of the aforementioned production channels. Concerning the s-channel process, there are predictions at NLO in QCD in terms of the total cross section [83] and fully differential cross sections [65], also including leptonic top-quark decays [64, 84]. NLO corrections for the t-channel process are also provided in terms of the total cross section [85–87] and fully differential cross sections [65, 88, 89], including leptonic decays of the top-quark [64, 90, 91]. There are also studies which compare five and four flavour scheme calculations of t-channel cross sections [62, 92]. They result in a reasonable agreement between the two approaches in case of observables which barely depend on the b-quark mass. Fully differential cross sections of associated Wt production are available at NLO in QCD as well [93, 94], also including leptonic top-quark decays [63]. Furthermore, there are combinations of fixed order results at NLO in QCD with parton showers regarding each of the three production modes in the context of Monte Carlo (MC) event generators [95–98]. A discussion of these generators is presented in Sec. 1.4.

In addition, effects of soft gluon radiation have been studied [71, 94, 99–103]. It can result in large logarithms near the kinematic threshold region, but it is possible to resum this contribution. However, results published in [100] could not be reproduced in the context of Ref. [70], which is documented in Ref. [104]. Accordingly, the LHC Top Working Group recommends to refer to NLO cross sections of s and t-channel single top-quark production and this is the choice made in this thesis [72]. Meanwhile also partial NNLO results on t-channel single top-quark production have been presented [105]. They are restricted to vertex corrections of two-loop amplitudes. Complete NNLO single top-quark cross section predictions remain to be finalised [106].

The number of available predictions indicates a thorough theoretical understanding of single top-quark production. In the following, total cross sections of s-channel single top-quark production are discussed in detail. This allows for a proper interpretation of the measurement presented in the last chapter of this thesis. All cross section calculations are performed using the HatHor program [70, 108]. It is dedicated to fast computations of total top-quark production cross sections at hadron colliders. For instance, a t-channel cross section computation at NLO with a statistical accuracy on the order of 10^{-3} performed with an Intel Xeon 2.4 GHz CPU takes about 100 ms using HatHor. Using MCFM 6.5 with default integration settings, on the other hand, it takes about 35 min. Clearly, HatHor allows one to perform comprehensive studies of total top-quark

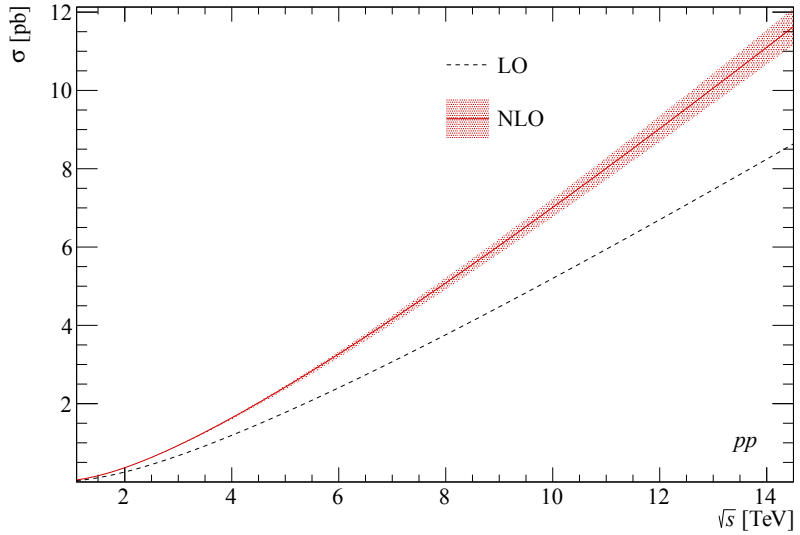


Figure 1.13.: The cross section of s-channel single top-quark production as a function of the centre-of-mass energy $\sqrt{s}$ in pp collisions. The calculations are performed at NLO in QCD using the CT10nlo PDF set [56]. The uncertainty bands indicate opposite variations of the scales μ_R and μ_F by factors of 2 and 1/2.

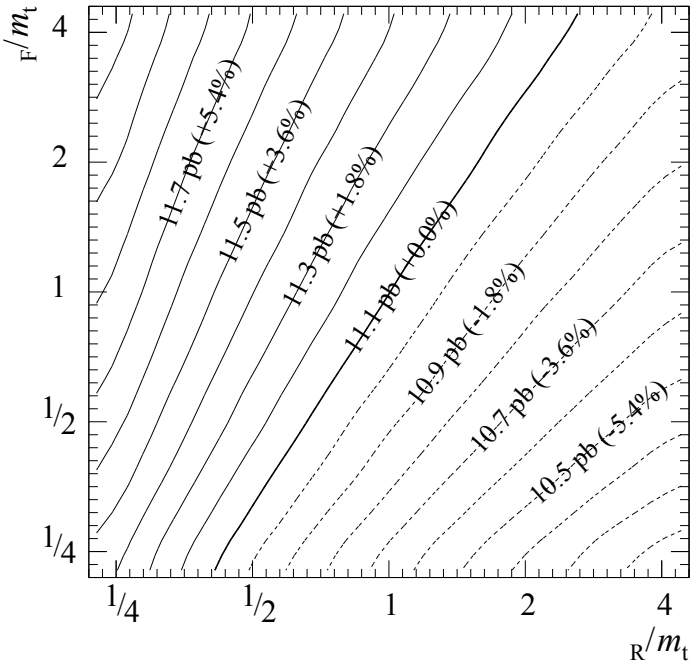


Figure 1.14.: Scale dependencies of the s-channel single top-quark production cross section in pp collisions at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$. The calculations are performed at NLO in QCD using the CT10nlo PDF set [56]. The bold contour line indicates the nominal cross section value at $\mu_R = \mu_F = m_t$. Due to correlations, the naive approach of simultaneous variations of both scales by equal factors would lead to an underestimation of the residual scale uncertainty [107].

production cross sections more efficiently. A description of the HatHor tool is given in App. B, accompanied by further studies of t-channel and associated Wt production cross sections.

Figure 1.13 shows the cross section of s-channel single top-quark production at the LHC as a function of the centre-of-mass energy $\sqrt{s}$, where variations of the scales μ_R and μ_F are taken into account. A more detailed study of scale dependencies is presented in Fig. 1.14, which shows s-channel cross section contours obtained from independent variations of the two scales. Deviations along diagonal contours $\mu = \mu_R = \mu_F$ are particularly small. Thus naive simultaneous variations of $\mu = \mu_R = \mu_F$ underestimate the scale uncertainty. For this reason Fig. 1.13 shows scale uncertainties obtained from opposite variations $(\mu_R, \mu_F) = (2 \cdot m_t, 1/2 \cdot m_t)$ and $(\mu_R, \mu_F) = (1/2 \cdot m_t, 2 \cdot m_t)$. Still, the scale variations are small, namely +3.3% and -2.7% at $\sqrt{s} = 8\,\text{TeV}$. However, the NLO corrections are quite large. The NNLO corrections would clearly be welcome.

The dependence of top-quark production cross sections on the top-quark mass is presented in Fig. 1.15. The single top-quark cross sections depend less on the top-quark mass than the top-quark pair production cross section. Experimental measurements of the top-quark mass meanwhile reached a high level of precision, $\Delta m_t^{\text{exp}} \lesssim 1\,\text{GeV}$, close to $\Lambda_{\text{QCD}} \approx 200\,\text{MeV}$. This gives rise to the question of how this parameter can be defined precisely. Since the top-quark is subject to strong interactions, it cannot be described in the same way as particles which can propagate freely like e.g. the electron. This issue is subject to ongoing research [111]. Anyway, the uncertainty of the s-channel cross section caused by the uncertainty of m_t is relatively small. The current experimental uncertainty given by the combination of Tevatron and LHC results, $m_t = 173.34 \pm 0.76\,\text{GeV}$ [7], translates into $\Delta\sigma^{m_t}_{s-\text{channel}} = \pm 1.7\%$. Due to its small impact, the uncertainty of the top-quark mass is neglected in the following.

Figure 1.16 presents PDF uncertainties of s-channel single top-quark production cross sections. Four different sets of PDFs are compared with each other, namely ABM11 [109], CT10 [56], MSTW08 [31] and NNPDF23 [58]. For each set, the uncertainties due to PDF fits and the strong coupling α_S are taken into account. All results are normalised to the cross sections calculated using the CT10 PDF set. The uncertainties of the individual PDF sets vary by about 2%. Except for the ABM11 PDFs, the different sets agree with each other to a relatively large extent. Thus PDF uncertainties on the s-channel cross section are small, which is expected since in this case the most important PDFs are those of the up and down-quark and their anti-particle counterparts. These PDFs are constrained well by lepton nucleon collision data.

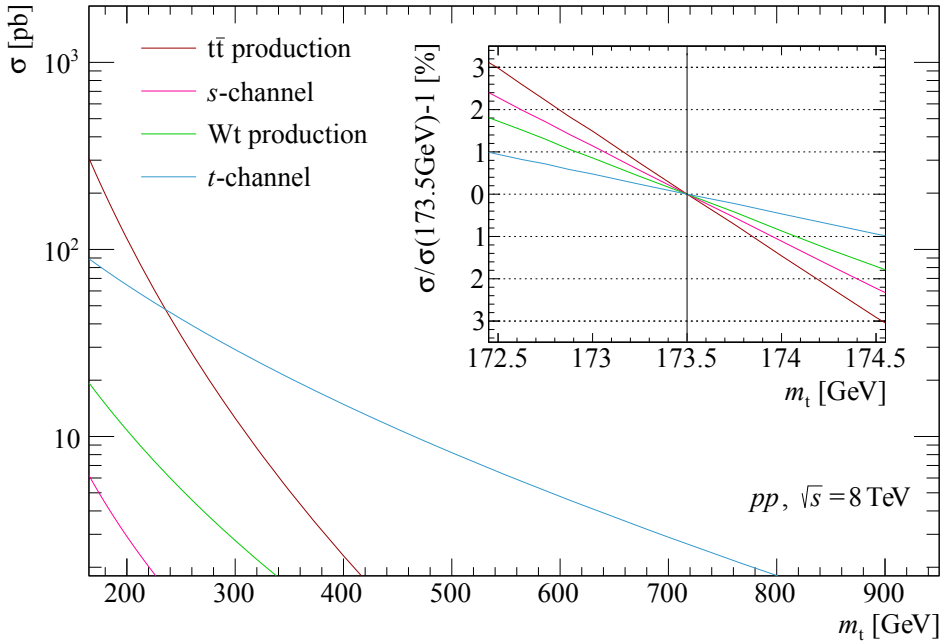


Figure 1.15.: Top-quark production cross sections in pp collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ as a function of the top-quark mass. The single top-quark production cross sections are calculated at NLO in QCD using the CT10nlo PDF set [56], whereas an NNLO calculation using CT10nnlo PDFs is performed in the case of the pair production [57]. The HatHor program uses the pole mass scheme for the top-quark mass. Top-quark pair production comes with the strongest mass dependence. However, in all cases the cross section uncertainty due to the uncertainty of the top-quark mass is at the per-cent level. In case of hypothetical heavier versions of the top-quark, single top-quark production is dominant [70].

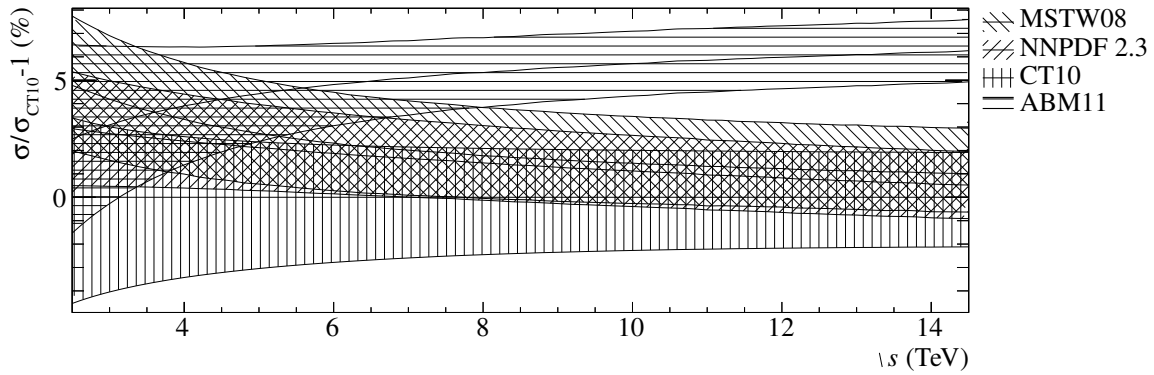


Figure 1.16.: PDF uncertainties of s-channel single top-quark production cross sections in pp collisions as a function of the centre-of-mass energy $\sqrt{s}$. Four sets of PDFs are compared with each other, namely ABM11 [109], CT10 [56], MSTW08 [31] and NNPDF23 [58]. The uncertainty bands are calculated according to the prescriptions of the PDF fitting groups, using Ref. [110] in the case of NNPDF23. All PDFs are employed at NLO accuracy. The predictions are normalised with respect to the CT10 results [107].

Collider	σ_s [pb]	σ_t [pb]	$\sigma_{t\bar{t}}$ [pb]	σ_t/σ_s	$\sigma_{t\bar{t}}/\sigma_s$
Tevatron $\sqrt{s}_{p\bar{p}} = 1.96\,\text{TeV}$	0.90	2.0	7.0	2.2	7.7
LHC $\sqrt{s}_{pp} = 8\,\text{TeV}$	5.2	85	240	16	46
LHC $\sqrt{s}_{pp} = 13\,\text{TeV}$	10	220	790	21	76

Table 1.2.: Top-quark production cross sections relevant to s-channel single top-quark searches at hadron colliders. The s and t-channel single top-quark production cross sections are calculated at NLO in QCD [62–70], while the $t\bar{t}$ cross section is calculated at NNLO in QCD [48–51, 66–69, 108]. A top-quark mass of $m_t = 173.3\,\text{GeV}$ and the MSTW2008NNLO PDF set are used [31]. The s-channel single top-quark production comes with the smallest rates, in particular at the LHC. This relation becomes even more distinct as the collision energy rises.

A common procedure to combine the predictions based on different PDF sets is described by the PDF4LHC working group [59]. The envelope of CT, MSTW and NNPDF cross section uncertainties due to PDF and α_S variations provides the overall uncertainty, while the midpoint of this envelope provides the central value. Clearly, this is not a rigorous but a pragmatic approach. However, in the present case the details of the combination of PDF sets are of minor importance due to the smallness of the deviations. In summary, the s-channel single top-quark production cross section at the LHC at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ is

$$\begin{aligned}\sigma_{s-\text{channel}}^{\text{LHC},\,8\,\text{TeV}} &= 5.24^{+0.15}_{-0.12}(\text{scales}) \pm 0.16\%(\text{PDF}+\alpha_\text{S})\,\text{pb} \\ &= 5.24^{+0.22}_{-0.20}\text{pb}\,. \end{aligned} \tag{1.39}$$

This is the cross section to which the measurement presented below in chapter 3 refers to.

1.3.5. Searches for s-channel single top-quark production

In the following a summary of past searches for s-channel single top-quark production is presented. The Tevatron collided protons and anti-protons at a centre-of-mass energy of $\sqrt{s} = 1.96\,\text{TeV}$. The increase of the amount of collision data and continuing improvements of measurement techniques allowed for separate observations of t-channel and finally in 2014 s-channel single top-quark production.

CDF uses a dataset corresponding to an integrated luminosity of up to $9.45\,\text{fb}^{-1}$ and selects collision events with a charged lepton and two jets requiring one or two b-tags. A complementary event selection involves missing transverse momentum and two or three jets with at least one b-tag. The s-channel signal is separated from the background processes by means of a machine learning technique, namely an artificial neural network. It considers kinematic variables and it is trained with the help of simulated events. The s-channel signal is identified with a significance of 4.2 standard deviations (s.d.) [112]. D0 uses a dataset corresponding to an integrated luminosity of $9.7\,\text{fb}^{-1}$ and selects events with 2 or 3 jets, at least one of which is b-tagged. Three different techniques are used in order to discriminate the s-channel signal from background processes. Aside from a neural network and boosted decision trees (BDTs), which constitutes another machine learning technique, a matrix element method is employed. The three methods perform similarly and result in correlations on the order of 75%. They are combined using another neural network, which leads to an identification of the signal with a significance of 3.7 s.d. [113]. Finally, s-channel single top-quark production is observed by means of a combination of these analyses, yielding a significance of 6.3 s.d. [75]. The measured s-channel cross section in $p\bar{p}$ collisions at a centre-of-mass energy of $\sqrt{s} = 1.96\,\text{TeV}$ is

$$\begin{aligned} \sigma_{s-\text{channel}}^{\text{Tevatron}} &= 1.29^{+0.26}_{-0.24}\,\text{pb} \\ &= 1.43^{+0.29}_{-0.27}\,\sigma_{s-\text{channel}}^{\text{SM, NLO}} \,. \end{aligned} \tag{1.40}$$

It agrees with the NLO SM expectation at the level of 1.5 s.d. (Tab. 1.2). The collaborations quote an expected total uncertainty of 20%, while the expected uncertainty in absence of systematic uncertainties is 14%. As the Tevatron has finished operations, this is most likely a final result. The CDF and D0 collaborations benefit from the Tevatron's anti-proton beam which provides large anti-quark PDFs leading to a relatively high s-channel rate. Thus the discrimination of the s-channel signal from its various backgrounds is easier than it is in case of the LHC which collides only protons. The cross sections of the s-channel signal and its two major backgrounds, namely top-quark pair production and t-channel single top-quark production, are compared with each other in case of the Tevatron and the LHC in Tab. 1.2.

At the LHC, CMS uses collision data at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ corresponding to an integrated luminosity of $19.3\,\text{fb}^{-1}$. Selected collision events contain one charged lepton and two or three jets, two of which are b-tagged. For each of the two jet multiplicities, BDTs are trained individually in order to discriminate signal and background contributions. No evidence for s-channel single top-quark production is found

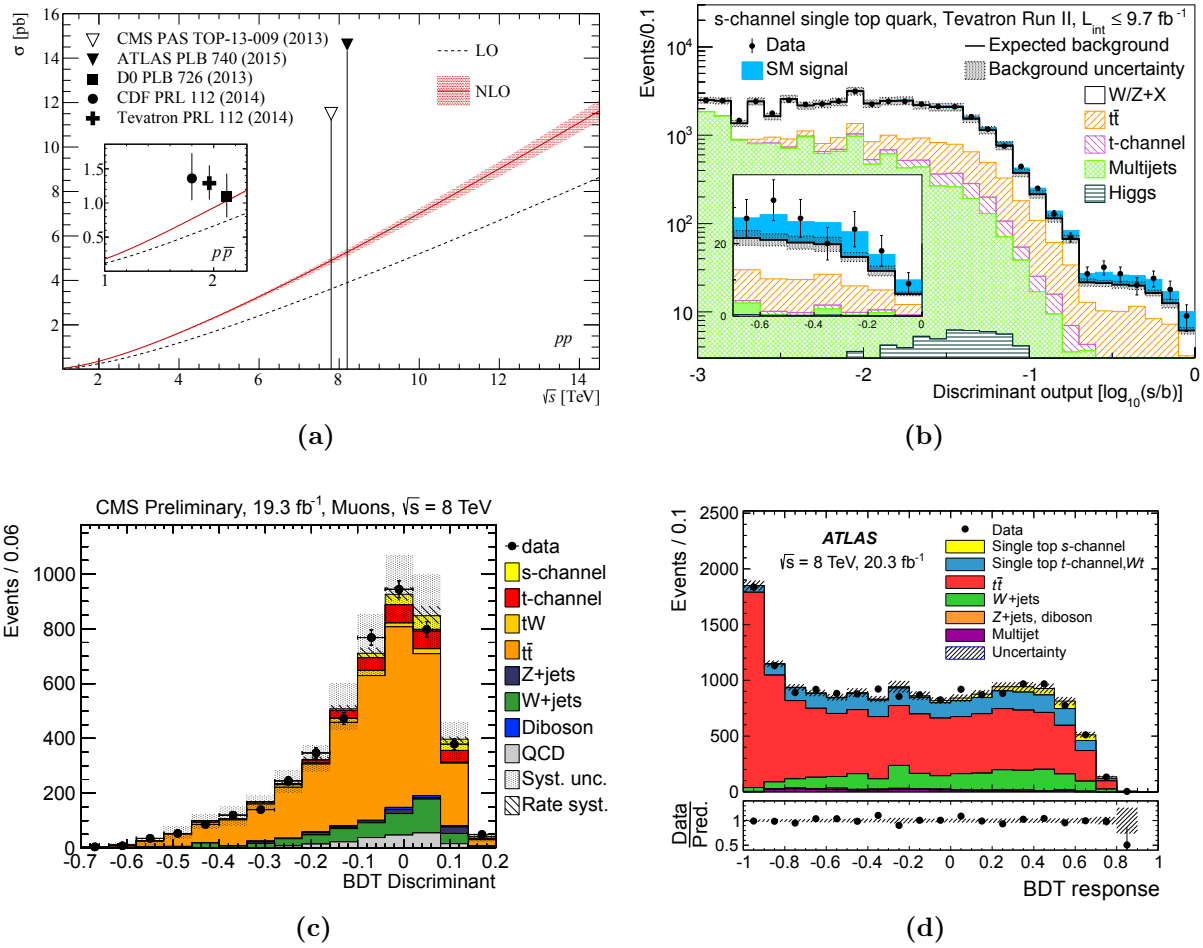


Figure 1.17.: Overview of searches for *s*-channel single top-quark production. In (a) experimental cross section measurements and upper limits are compared with theoretical predictions at NLO accuracy, where the error bands indicate scale uncertainties. The Tevatron combination (b) leads to the observation of the process. At the LHC the signal is accompanied by higher background rates, as seen in the distributions of BDT responses by CMS (c) and ATLAS (d) [75, 112–116].

as the determined signal significance amounts to only 0.7 s.d. An upper limit of 11.5 pb is set on the production cross section at a confidence level (CL) of 95% [115]. ATLAS analyses *pp* collisions at the same centre-of-mass energy with an integrated luminosity of 20.3 fb^{-1}. Events with one charged lepton, missing transverse momentum and two *b*-tagged jets are selected. BDTs are used in order to discriminate the *s*-channel signal from its backgrounds. Again, no evidence for *s*-channel single top-quark production is found as the signal can only be identified with a significance of 1.3 s.d. An upper limit of 14.6 pb is set on the production cross section at 95% CL [116]. An overview of *s*-channel discriminant distributions obtained by the various measurements is presented in Fig. 1.17.

In summary, it is fair to say that s-channel searches are challenging, in particular at the LHC.

1.4. Simulation of hadronic collision events

The interpretation of collision data is a demanding task, and it is particularly difficult in case of hadronic collisions due to the complex phenomenology of the strong interactions. Only in special cases it is possible to restrict measurements to simple observables, an example being the invariant mass of two photons in Higgs boson measurements, $pp \to H^0 + X \to \gamma\gamma + X$. In almost all other cases a detailed understanding of the kinematics of collision events is mandatory. This includes the complex final states of top-quark events which need to be understood in detail in order to disentangle signal and background contributions. Several different aspects are involved in this task, from hard scattering processes through to responses of detector electronics. Accordingly, there is no complete analytic approach to it, but rather simulations of consecutive steps are performed. First of all, the collision event itself is simulated up to the level of final state hadrons, leptons and photons. In a second step, the detector response to these particles is simulated. While the latter part is described below in Sec. 2.4, the simulation of the primary scattering is described in the following. The discussion is restricted to a brief overview of this task, which is subject to ongoing research [111].

1.4.1. Factorization of scattering processes

The scattering process involves different energy scales and its simulation is factorized accordingly. The hard scattering process which involves the highest scale is simulated first. Subsequently, additional partons radiated off of partons involved in the hard scattering are generated by means of the parton-shower approximation. In addition, multiple parton interactions occurring in the same hadronic collision are generated, known as the underlying event, which mostly adds low p_{T}, or soft particles. Then the partonic final state evolves into a set of hadrons, some of which are short-lived, so that their decays are simulated as well. In addition, soft interactions of other colliding particles which accompany the primary collision in the crossing of particle beams are added to the event. Figure 1.18 illustrates the factorized simulation of collision events. In the following, major aspects of this approach are discussed.

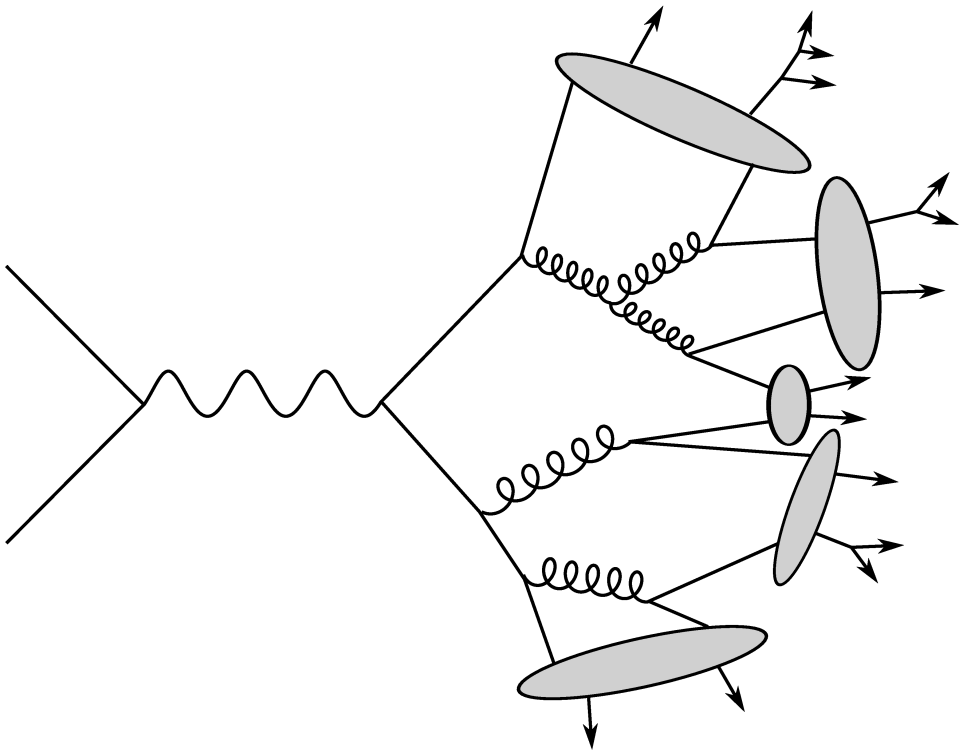


Figure 1.18.: Collision event simulation. The hard scattering, which involves the highest energy scale is simulated first. It is illustrated in the left-hand part of the figure. Subsequently, a parton shower creates additional partons at intermediate scales. For simplicity, only a final state parton shower is shown here, leaving out the initial state. The scale of parton splittings within the shower decreases step by step. Once a low scale on the order of 1 GeV is reached, hadronization takes place, as illustrated by the grey areas. Here partons are combined into hadrons, some of which decay quickly.

Hard scattering

This simulation step is basically a cross section computation. As the phase spaces of final states of interest involve many dimensions in general, an efficient integration procedure is needed. Therefore MC sampling techniques are used, since their rate of convergence does not depend on the dimensionality of the integral. During the sampling of the phase space, individual points are gathered in order to obtain a set of collision events. While it is in principle always possible to generate sets of unweighted events, some event generators assign a weight to every event, which allows them to scan the phase space more efficiently. Anyway, sets of complete parton-level events are obtained as a result, aside from the total cross section. In the leading order of perturbation theory this procedure can be performed straightforwardly. Meanwhile some event generators even provide NLO accuracy. Their combination with parton showers is rather involved. In any case,

the particle multiplicities which can be taken into account in the hard scattering are relatively low. In order to reach desired multiplicities, the parton shower approximation takes over.

Parton shower

The aim of parton showers is to enrich the final states of hard scatterings with additional partons. Two approaches are available. In the case of dipole showers, partons are emitted by pairs of other partons. A more conventional but still accurate approach is based on successive emissions by single partons. Here partons generated in the hard scattering are assigned to an increased mass in order to allow for parton splittings. Energy-momentum conservation is restored as soon as the showering has ended. Starting from the assigned high mass Q, typically on the order of the hard scattering scale, an evolution down to a non-perturbative scale Q_0 is performed by means of parton splittings. At each splitting, the outgoing partons are assigned to a lower scale until Q_0 is reached and the showering stops. The scale Q_0 defines a resolution criterion, so that splittings are taken into account only if they can be resolved, for instance if the transverse momentum of the radiated parton relative to its parent parton is sufficiently large.

Even though this procedure seems classical in the first place, it actually is derived from QCD in the limit of small angle, or collinear radiation. In fact, in the collinear limit the cross sections of arbitrary scattering processes in QCD where partons of type i are accompanied by a parton of type j factorize into the lower order contribution σ_0 and universal splitting terms,

$$\mathrm{d}\sigma \approx \sigma_0 \sum_{\{\text{partons } i\}} \frac{\alpha_S}{2\pi} \frac{\mathrm{d}\vartheta^2}{\vartheta^2} \mathrm{d}z \mathrm{d}\varphi P_{ji}(z, \varphi)\,. \tag{1.41}$$

Here z denotes the momentum fraction of the radiated parton with respect to its emitter, while ϑ and φ provide its direction. In particular, the opening angle of the splitting, ϑ, can serve as an evolution variable of the shower. The sum over partons involved in Eq. 1.41 allows one to perform splittings of individual partons. Using the summand, the distribution of the hardest emission can be derived. It is generated by means of MC techniques. If the resulting scale ϑ is greater than the threshold ϑ_0 defining resolvable radiations, a corresponding parton is added. This procedure is repeated with respect to all final state partons including the ones resulting from the shower, until none of the generated emissions are resolved. While evolution variables other than ϑ can be used

in principle, angular ordered showers turn out to also take soft emissions into account correctly. Furthermore, there are also showers of initial states. They could be performed similarly, but it is more efficient to reverse the showering sequence, generating partons of increasing momenta.

Effectively logarithms which appear in the collinear limit are summed to all orders by parton showers. While they explicitly deal with the probability of resolved radiation to occur, they also include unresolved radiation since the two probabilities sum up to one. In other words, loop corrections are taken into account implicitly. Still, NLO precision is not reached as the procedure is restricted to the collinear limit. Also the total cross section remains unchanged. Meanwhile event generators are available which achieve a merging of hard scatterings and parton showers at NLO. However, the details of hard scattering and parton shower matching and merging are beyond the scope of this overview [111, 117–120].

Underlying event

Apart from partonic interactions which result in high-p_{T} objects triggering experimental event selections, further interactions can occur within the same hadronic collision. These multiple parton interactions (MPIs) can even lead to additional jets. However, they mostly produce low-p_{T} particles. This additional contribution must be taken into account in order to describe experimental data correctly, as shown in Fig. 1.19. Frequently used event generators make use of the Sjöstrand-Zijl model, which describes MPI in an essentially non-perturbative way [121]. The model is rather simple. In particular, correlations in terms of colour among MPI systems are relatively poorly understood to date. In fact, recent research indicates some shortcomings of current MPI models [122].

Hadronization

The transition from partonic to hadronic final states is denoted by hadronization. No rigorous approach to a modelling of this process has been developed to date. Instead, event generators employ models inspired by QCD which involve several tunable parameters. In the string model, popular in form of the Lund model [124, 125], a linear rise of potential energy at large distances between coloured particles is assumed. If quarks are separated from each other, the growth of potential energy eventually leads to the creation of new quarks out of the vacuum, with which the original quarks combine and form hadrons.

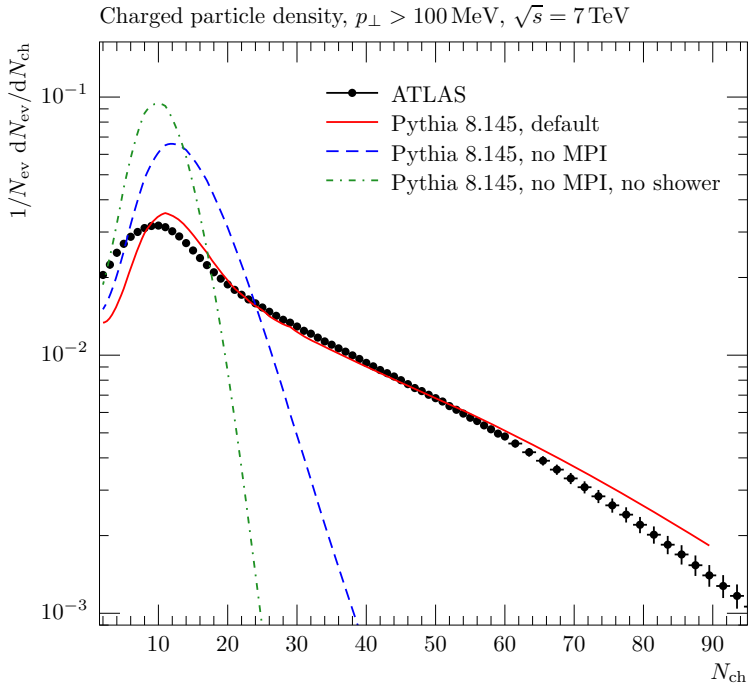


Figure 1.19.: Charged particle multiplicity in pp collisions as measured by the ATLAS collaboration [123]. The measurement is compared with different predictions provided by the Pythia 8 event generator. Clearly, MPI and parton showers must be taken into account in order to describe the data correctly [111].

The cluster model, on the other hand, is based on a property of parton showers called preconfinement. Colour singlet combinations of partons can be formed with an invariant mass distribution which is independent of the scale of the hard scattering process [126]. Typical cluster masses are on the order of 1 GeV, which is the hadronic mass scale. The clusters decay, which results in hadrons. Subsequently, decays of short-lived hadrons are simulated. In addition, also QED radiation as well as τ lepton decays are considered. After these steps, the generated final state can finally be passed on to a detector simulation.

1.4.2. Selected event generators

Several event generators are available. They differ with respect to their simulation scope, perturbative order, matching and merging of parton showers and hard scatterings as well as the employed models used beyond the hard scattering. Furthermore, regarding NLO precision generators, two different schemes have been developed, known as the Powheg and MC@NLO schemes. In practice, different generators are used for different purposes

Generator	Precision	Beyond $\int \lvert\mathcal{M}\rvert^2$	Most Common Usage
AcerMC	LO	–	Hard Scattering
Alpgen	LO	–	Hard Scattering
(a)MC@NLO	NLO	–	Hard Scattering
Herwig	NLO	✓	PS/UE/Had/...
Powheg	NLO	–	Hard Scattering
Pythia	LO	✓	PS/UE/Had/...
Sherpa	NLO	✓	All

Table 1.3.: Event generators used in this thesis. More and more generators can simulate hard scatterings at NLO. Herwig, Pythia and Sherpa allow one to model processes beyond the hard scattering, including parton showers (PS), the underlying event (UE) and hadronization (Had). The two former ones are often combined with other hard scattering generators, while Sherpa is constructed to be an all-in-one generator. More detailed information on the employment of the generators in the measurement presented in this thesis follows in chapter 3 [117, 120, 127–134].

on a case by case basis, depending on the goodness-of-fit to experimental data. Table 1.3 provides an overview of generators used in the measurement presented in this thesis.

Chapter 2.

Experimental setup

"Ich warte auf die dunklen Massen
I'm waiting for the dark masses
zwischen den Sternen noch unentdeckt
between the stars still undiscovered"

— Einstürzende Neubauten, *Ich warte*

2.1. The Large Hadron Collider

The LHC is a circular proton collider of 27 km circumference built at CERN near Geneva. It is located about 100 m underground, crossing the border of Switzerland and France. Aside from protons, it can also collide lead ions. However, this thesis will only deal with proton-proton collisions, which is also the focus of the LHC design and operation. Collisions are delivered to four large experiments, namely ATLAS and CMS, which are multi-purpose detectors, LHCb, which is specialised in analysing B-meson decays, and ALICE, which is designed for studies of quarks and gluons at high density in lead ion collisions.

Built in the tunnel of its predecessor, the Large Electron-Positron Collider LEP [135], it is a part of CERN's accelerator complex. While the e^+e^--collisions in LEP were limited to a centre-of-mass energy of $\sqrt{s} = 209\,\text{GeV}$ due to the large amount of synchrotron radiation emitted by the light electrons and positrons, the LHC was designed for pp-collisions at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$. Due to the higher mass of its colliding particles, synchrotron radiation is a minor concern for the LHC. On the other hand, the momentum fractions of the colliding proton constituents are not known and the collisions lead to

complex final states as outlined in chapter 1. Furthermore, most of the physics analyses need to deal with large backgrounds caused by strong interactions. Therefore the LHC is mostly considered to be a discovery machine rather than a precision measurement device. In terms of particle beams, it provides the highest collision energies reached in an artificial particle accelerator built to date, while the collision rates are unprecedented regarding hadron colliders. The operation of the LHC initially lasted from 2008 until early 2013 and was continued in 2015 after a first long shutdown required for maintenance work.

In this section the design of the LHC is briefly described. Here the ultimate operation parameters of the LHC are quoted while the actual operation parameters used in 2012, relevant to the measurement presented in this thesis, are discussed in Sec. 2.3. More detailed information about the LHC can be found in Ref. [136, 137].

2.1.1. CERN's accelerator complex

The LHC is supplied with proton beams by a couple of pre-accelerators, which are connected in series (Fig. 2.1). They were built for the purposes of previous experiments over decades and have been upgraded in order to meet the needs of the LHC. First of all, protons are obtained from a duoplasmatron. Here a cathode filament emits electrons which subsequently ionize a hydrogen gas. A beam of the resulting free protons is formed and accelerated by the linear accelerator LINAC 2. With an energy of 50 MeV this beam enters the Proton Synchrotron Booster (PSB). In the four rings of the PSB, the beam is accelerated up to an energy of 1.4 GeV before it is injected into the Proton Synchrotron (PS). The bunch structure of the beam is enhanced further by the PS, creating patterns of bunches displaced by a constant offset. In 2012 this offset amounted to 50 ns while the design value is 25 ns. The creation of bunches is necessary in order to allow for an acceleration of the beam by means of radiofrequency cavities. In addition, the detectors which record collisions later on need to be read out at discrete points in time. Between so-called bunch trains some extra space is left empty in order to allow for proper beam transports by ejection kicker magnets. Once the PS accelerated a bunch train to an energy of 25 GeV, it is ejected from the PS and guided to the Super Proton Synchrotron (SPS). The SPS accelerates the beam until an energy of 450 GeV is reached, before the beam injection into the LHC can start. For this purpose, the SPS uses two different transfer lines, thereby splitting the single beam into two distinct beams, one of which

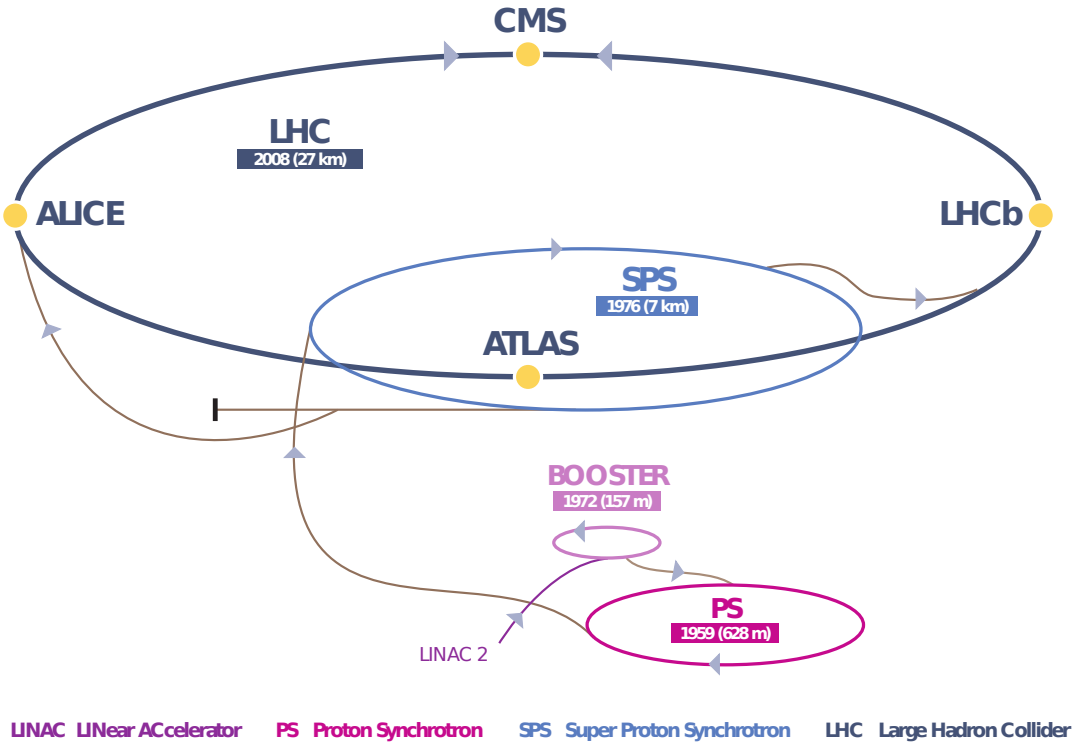


Figure 2.1.: CERN's accelerator complex. Only components relevant for the *pp*-operation of the LHC are shown. Several pre-accelerators are used to provide two proton beams of an appropriate structure and energy to the LHC. The LHC itself provides collisions to the four experiments ALICE, ATLAS, CMS and LHCb [138].

circulates clockwise while the other one circulates counter-clockwise in the LHC. During this injection, the bunch train structure is conserved for each of the two ejected beams.

The LHC inherited eight arcs and eight straight sections from the LEP geometry. Each of the straight sections is about 500 m long and can be used for purposes of collision experiments or beam setup. ATLAS is located close to the SPS while CMS is located on the opposite side of the LHC. With respect to the beam rotating clockwise, ALICE is located in the first octant following ATLAS while LHCb is located in the last octant before ATLAS. In the octant before CMS, two independent sets of eight superconducting radiofrequency cavities are installed. They accelerate the two beams up to their final energy. Each of the cavities provides a voltage of 2 MV corresponding to a gradient of 5.5 MV/m. This leads to an increase in energy of 485 keV per turn. The dumping of the beams is accomplished in the first octant following CMS.

2.1.2. Magnets

Since high magnetic field strengths are needed at the LHC, finite resistance conductors cannot be employed as they would produce too much heat given the necessary currents. Therefore LHC magnets are made of superconductors. Superconducting NbTi filaments form strands which are combined into cables. These configurations, called Rutherford cables, provide a superconductor with a large surface, which is where the current actually flows, without reaching the critical values of the current density and field strength. The Rutherford cables are arranged in a way which leads to the desired magnetic field topologies. All superconductors are cooled down to an operating temperature of 1.9 K using superfluid helium. Space limitations in the tunnel as well as financial aspects led to a "two-in-one" design for almost all of the superconducting LHC magnets. Here both beams circulate in one magnet device sharing the same cryostat.

The LHC uses several types of magnets to deflect, focus and transfer the beams. For the purpose of deflection, dipole magnets are used. They are designed to operate with a current of 12 kA, leading to a magnetic field of 8.3 T. This is equivalent to the design proton beam energy of 7 TeV. 1232 of these main dipoles are installed. Quadrupole magnets operated at the same current with a design gradient of 223 T/m are used to focus the beams. Additional higher multipole magnets serve for corrections of the beam optics. Close to the interaction regions, magnets compiled by dedicated quadrupoles squeeze the beams in order to provide high luminosity through small transverse beam sizes at the interaction points (IPs).

2.1.3. Proton beams

The rate of a scattering process depends on its intrinsic strength, given by the cross section σ_{proc}, and on a set of beam parameters, summarised by the luminosity $\mathcal{L}$:

$$\frac{\mathrm{d}N_{\text{proc}}}{\mathrm{d}t} = \mathcal{L} \cdot \sigma_{\text{proc}} \ . \tag{2.1}$$

Accordingly, the aim of the LHC operation is to provide high luminosities in order to extend the physics potential of the experiments as far as possible. If a Gaussian particle

density profile of the bunches is assumed, the luminosity can be expressed as follows:

$$\mathcal{L} = \frac{n_b N_{p/b}^2 \nu_{\mathrm{rev}} \gamma}{4\pi \varepsilon_n \beta^*} F(\theta_c) \ . \tag{2.2}$$

Here n_b denotes the number of bunches per beam while $N_{p/b}$ denotes the number of particles per bunch. ν_{rev} denotes the revolution frequency. It is fixed by the LHC circumference and by the speed of light. $\gamma = 1/\sqrt{1-(v/c)^2}$ is the relativistic gamma-factor. ε_n denotes the transverse beam emittance, while β^* is the value of the beta-function at the IP. The beam particles undergo transverse oscillations around the nominal orbit with an amplitude of $\sqrt{\varepsilon_n \beta}$ in each of the two directions. While the emittance stays approximately constant, the β-function is squeezed down to small values at the IPs. For the LHC also a geometry factor $F(\theta_c)$ enters because the beams collide under a finite angle θ_c. This setup is necessary because head-on collisions would give rise to multiple IPs along the experiments since the distance between bunches is smaller than the sizes of the detectors. The most relevant beam parameters of the LHC are given in Tab. 2.1. They result in a design peak luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Besides the high proton energy, it is this high luminosity which establishes the importance of the LHC. Several effects limit the luminosity. In particular these are beam-beam interactions experienced by the protons during collisions and the peak β-function value combined with the mechanical aperture of the beam pipe. The maximum relativistic gamma factor is limited by the dipole magnets with their field strength of 8.3 T.

Once the LHC is filled with proton beams, has ramped the beam energy to the desired value and has squeezed and crossed the beams at the IPs, collisions occur at a peak luminosity. In the course of the following operation, the luminosity decreases due to the collisions at the IPs, proton collisions within bunches causing the emittance to grow and beam interactions with residual gas in the vacuum chamber. The luminosity decrease can be approximately described by an exponential decay. Typical luminosity lifetimes are on the order of 15 h.

Number of bunches per beam n_b	2808
Number of protons per bunch $N_{p/b}$	$1.15 \cdot 10^{11}$
Revolution frequency ν_{rev}	11.2 kHz
Relativistic gamma	7461
Transverse beam size	16.7 µm
Geometric luminosity reduction factor F	0.836

Table 2.1.: LHC design beam parameters. The transverse beam size and geometric luminosity reduction factor F refer to the ATLAS and CMS IPs.

2.2. The ATLAS detector

ATLAS (*A Toroidal LHC Apparatus*) is a multi-purpose detector used to record the collisions at the corresponding interaction point of the LHC. Its design is driven by the aim of analysing the physics of electroweak symmetry breaking and possible new phenomena at a similar scale in pp collisions. For this purpose ATLAS is broadly equipped. It has to meet many needs dictated by the dense LHC environment and by its physics goals. A large geometrical coverage and radiation-hard electronics and sensors are needed. High granularity detector components are required in order to reconstruct final states with many particles coming from the vertices of the scattering process of interest and from the additional pile-up interactions occuring in the same bunch crossing. Many different physics objects must be distinguished. In particular, an efficient and pure reconstruction of high p_{T} electrons, muons and photons is needed because the presence of these particles indicates the occurence of many of the most interesting electroweak processes. The reconstruction of physics objects must be realized with high resolutions in order to allow for the identification of the interesting primary heavy particles decaying already within the beam pipe. Finally the acquisition of the data must be performed in an efficient manner with a fast detector readout taking into account the distance between LHC proton bunches of 25 ns. ATLAS accomplishes these goals.

Figure 2.2 shows an overview of the ATLAS detector. It is symmetric with respect to rotations about the beam axis and reflections with respect to the transverse plane at the IP. The overall apparatus is compiled by inner tracking detectors, electromagnetic and hadronic calorimeters and a muon spectrometer. The inner part of the central barrel region contains a solenoid magnet while the muon spectrometer contains toroid magnets in the barrel and forward regions. The muon spectrometer defines the size of the

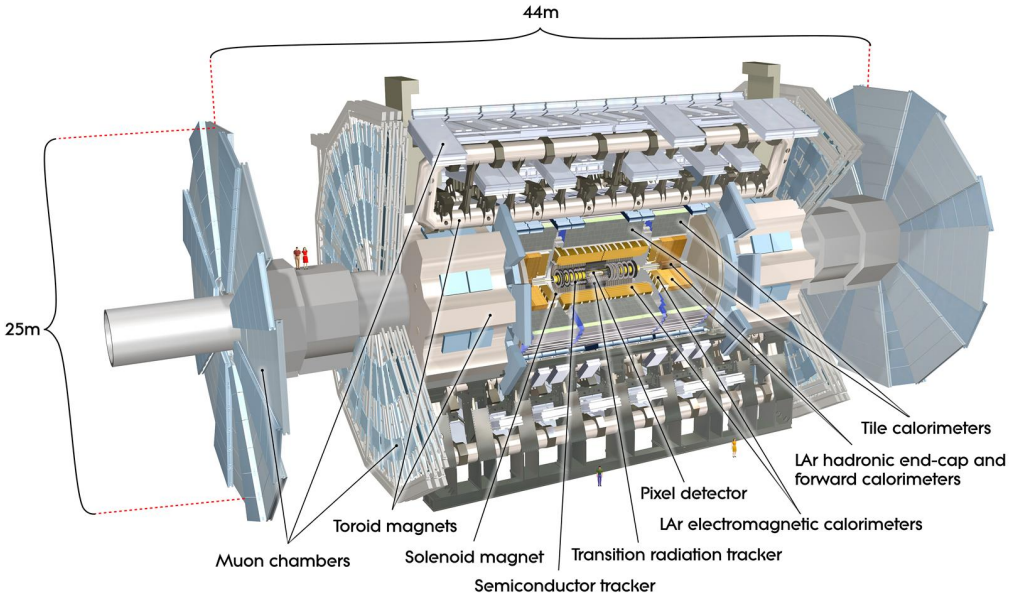


Figure 2.2.: The ATLAS detector. The IP is surrounded symmetrically by different layers of detector components which allow for a detailed reconstruction of collision events [139].

ATLAS detector with a length of 44 m and a diameter of 25 m. Altogether the detector weighs 7000 t. An overview of the ATLAS site with caverns containing the detector and supporting equipment as well as surface infrastructure buildings is given in Fig. 2.3.

In the following, after a discussion of the employed coordinate system, a short description of the components of ATLAS is given. More detailed descriptions can be found in Ref. [3, 140, 141].

Coordinate system

The ATLAS coordinate system used throughout this thesis is right-handed and the primary interaction vertex serves as the origin. The x-axis points to the centre of the LHC ring, while the y-axis points to the surface of the earth. Accordingly, the z-axis is aligned with the direction of the beam rotating counter-clockwise. The azimuth angle φ is measured in the x-y-plane against the x-axis while the polar angle ϑ is measured against the z-axis. It is a common practice to use the rapidity y instead of the polar angle ϑ and, as most of the final state particles are approximately massless, the pseudorapidity

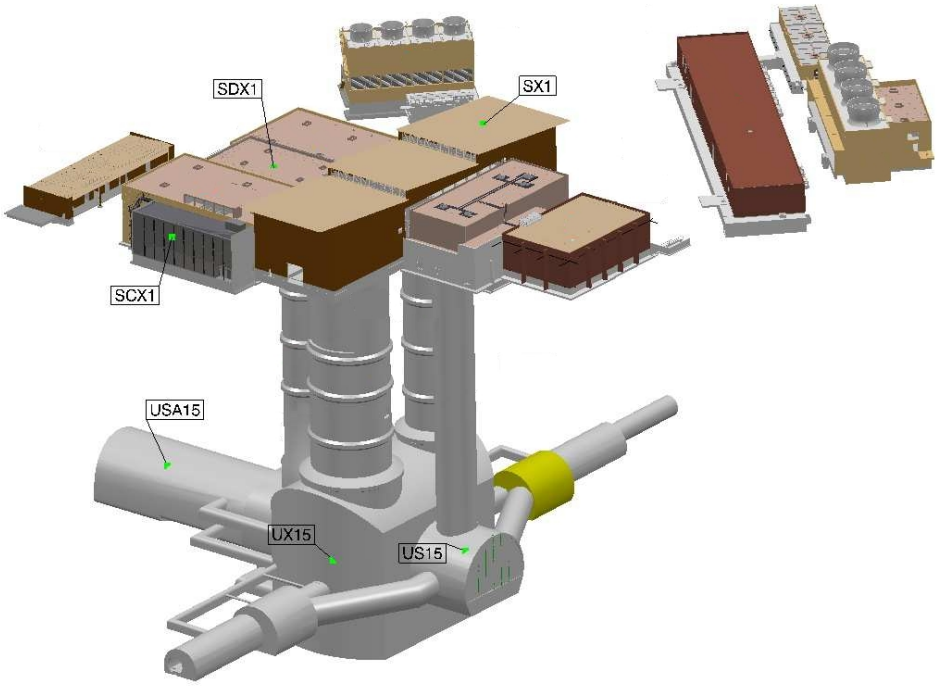
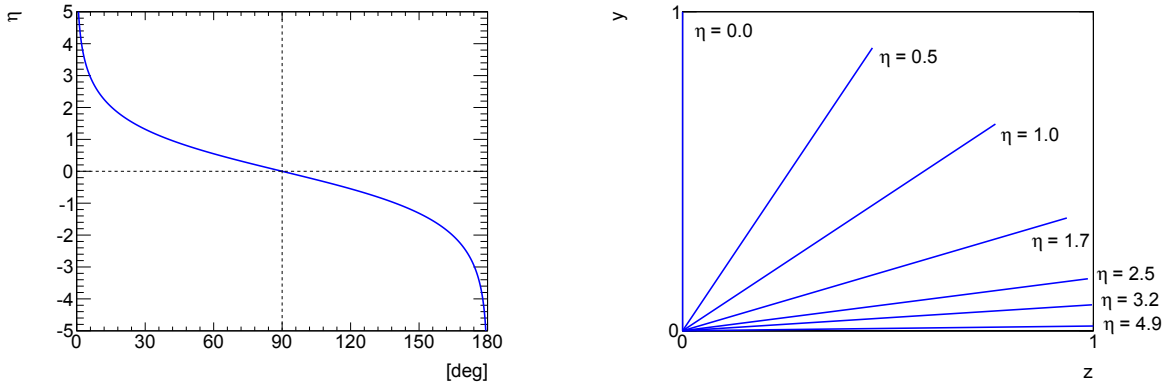


Figure 2.3.: The ATLAS infrastructure. The main cavern UX15 houses the ATLAS detector. It is surrounded by the counting rooms USA15 and US15 which house the back-end electronics. In USA15 there is also service equipment which needs to be located close to the detector, e.g. cooling systems. The underground levels are connected to the surface by several shafts. The main shaft connects the detector cavern with the SX1 building which houses cranes. Here detector components were lowered into the main cavern. The SDX1 building is located on top of the personnel access shaft and houses the data acquisition room among others. The ATLAS control room is located in the SCX1 building. In addition, there are several supporting buildings, e.g. cooling towers dissipating heat recovered from the site via water cooling circuits [140].



(a) **(b)**

Figure 2.4.: Illustration of the pseudorapidity. In (a) the dependence of the pseudorapidity η on the polar angle ϑ is shown. Some examples of directions with their corresponding pseudorapidity are shown in (b). Here the higher η values, starting from $\eta = 1.7$, show the coverage of ATLAS detector components discussed below, namely the hadronic tile calorimeter, the inner tracking detectors, the end-cap calorimeters and the forward calorimeters.

η is an equivalent choice,

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \ ,$$
$$\eta = -\ln \tan \frac{\vartheta}{2} \ , \quad \eta = y \ \text{ if } \ m = 0 \ .$$

The usage of rapidity is motivated by its transformation under Lorentz boosts along the z-axis. Here η is only changed by a constant offset depending on the velocity of the boost. Accordingly, differences of pseudorapidities are invariant under these boosts. This property is important, as the momentum fractions of colliding partons and thus also the partonic centre-of-mass frame is unknown. However, differences of pseudorapidities $\Delta\eta_{ij} = \eta_i - \eta_j$ are not affected by this fact. Consequently, differences of particle directions are expressed by

$$\Delta R_{ij} = \sqrt{(\Delta\eta_{ij})^2 + (\Delta\varphi_{ij})^2} \ . \tag{2.3}$$

Another motivation of the usage of rapidities is its natural appearance in the phase space measure $\mathrm{d}^3p \sim \mathrm{d}y\mathrm{d}{p_\mathrm{T}}^2\mathrm{d}\varphi$. Many spatial distributions of observables at hadron colliders are relatively flat with respect to pseudorapidity. An illustration of the pseudorapidity is provided in Fig. 2.4.

2.2.1. Magnets

Four magnets enforce the deflection of charged particles in ATLAS. First there is a central solenoid (CS) in the inner barrel part between the tracking detectors and the electromagnetic calorimeter. Downstream of the calorimeters there is a barrel toroid (BT) in the central part, while two end-cap toroids (ECTs) are installed in the forward direction on each side of ATLAS. The ECTs are inserted into the BT lining up with the CS. NbTi superconductors packed in Rutherford cables are used in each of the magnets. They are cooled down to about 4.5 K using liquid helium. The energy stored in the ATLAS magnets amounts to 1.6 GJ which e.g. surpasses the energy of 0.6 GJ stored in the LHC magnets.

The CS provides a 2 T axial magnetic field for the inner tracking detectors. Regarding the performance of the following electromagnetic calorimeter, the material budget of the CS needed to be minimized. Therefore it shares one cryostat with this calorimeter. The BT and ECTs provide toroidal magnetic fields for the muon spectrometer with field strengths of 0.5 T and 1 T in the central and forward regions, respectively. Each of these toroid magnets consists of eight racetrack-shaped coils mounted radially around the z-axis. The BT contains one cryostat per coil while there is one large cryostat for each of the ECTs.

2.2.2. Inner tracking detectors

The ATLAS inner tracking detectors serve the purposes of vertex reconstruction, charged particle momentum measurement and charge discrimination. Furthermore, transition radiation is used for the identification of electrons. The overall tracking system is contained in a cylindrical envelope supported by the central solenoid magnet. It consists of three components, the granularity of which decrease with their distance from the beam pipe. Starting from the innermost layer, these components are a semiconductor pixel tracker, a semiconductor strip tracker (SCT) and a transition radiation tracker (TRT). The semiconductor part extends over a range of $|\eta| < 2.5$, while the TRT extends over a range of $|\eta| < 2.0$. Each of the three components is divided into concentric cylinders in the barrel region and end-cap disks perpendicular to the beam axis in the forward regions (Fig. 2.5, 2.6). Being subject to large irradiation doses, all components were designed to allow for an operation over a period of ten years at LHC design luminosity except for the

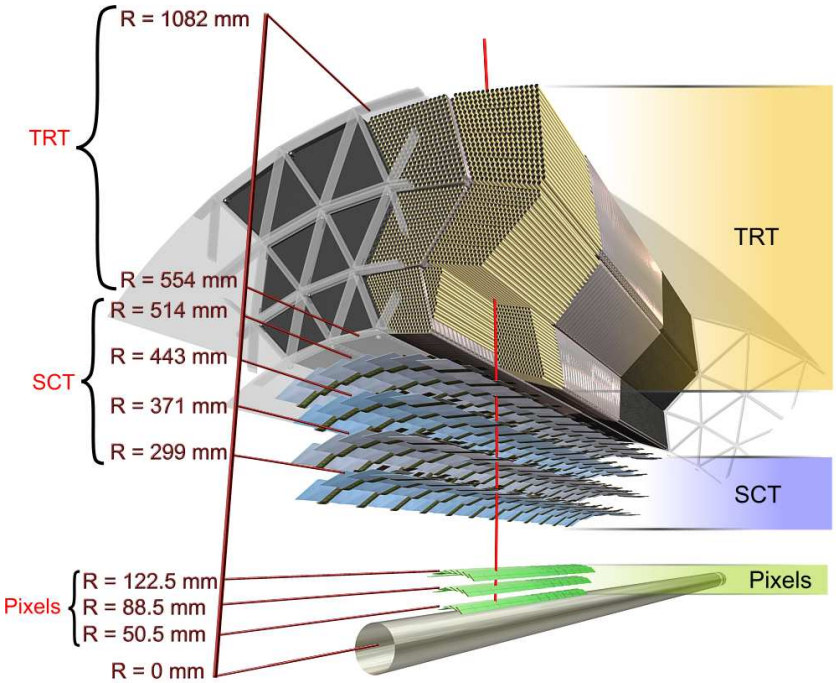


Figure 2.5.: ATLAS inner tracking detectors, barrel region. The red line shows a central charged particle track of $p_{\mathrm{T}} = 10\,\mathrm{GeV}$ traversing the beam pipe, the pixel detector, the SCT and the TRT [141].

innermost pixel layer. The pixel and SCT detectors are connected to a cooling system which extracts their heat and cools them down to about $-5\,^{\circ}\mathrm{C}$ to $-10\,^{\circ}\mathrm{C}$ in order to keep down noise after irradiation. The inner tracker alignment is monitored with the help of measured tracks by minimising the residuals between the positions of module hits and the fitted track position at each module. Additionally, the SCT alignment is supported by an interferometric system. The overall momentum resolution provided by the inner tracker is approximately given by

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = \sqrt{(0.05\%\, p_{\mathrm{T}}[\mathrm{GeV}])^2 + (1\%)^2}\,. \tag{2.4}$$

In the following a short overview of the setup of each of the three subdetectors is given.

Pixel tracker

The pixel tracker is the first detector component installed next to the beam pipe. In particular it serves the reconstruction of track impact parameters and secondary vertices, caused e.g. by decays of B-mesons or τ-leptons. Therefore, it is the component with the highest granularity of the ATLAS detector. The three barrel layers and six end-cap disks

consist of modules, while each module contains a sensor which in turn is connected to readout chips. These sensor tiles are made up of 250 µm thick n-doped silicon connected to a p-doped backplate. Each of them contains 47 532 n^+-doped pixel implants of a size of about 400 µm × 50 µm. With each pixel being bump-bonded to the front-end readout of its module, there are about 80 million pixel readout channels. The usage of n^+ implants instead of simple pn-junctions allows for detector operation even after the occurrence of semiconductor type inversion caused by high irradiation doses. The pixel modules are depleted by a bias voltage between 150 V and 600 V depending on the degree of radiation damage. The collection of ionisation charges and the signal readout take less time than the distance between LHC bunch crossings of 25 ns. Due to their high granularity, the pixel channels have a low occupancy on the order of 10^{-4}. As the innermost barrel layer of the pixel detector is exposed to the largest irradiation doses, a new innermost layer was inserted into ATLAS. This insertable B-layer has been installed in 2014.

Semiconductor strip tracker

The SCT follows the pixel detector. As it is located further away from the IP than the latter, its reduced granularity is sufficient. The four barrel layers and 18 end-cap disks of the SCT contain sensor modules. Each module consists of two strip sensors mounted back-to-back on a common baseboard with a stereo angle of 40 mrad. The SCT sensors are single-sided p-in-n silicon microstrip detectors. The n-doped bulk is 285 µm thick and contains 768 readout strips with a pitch between 57 µm and 90 µm depending on the position of the module. The bias voltage applied varies between 150 V and 350 V depending on the degree of radiation damage. Each strip is wire-bonded to the readout electronics giving about 6 million readout channels in total. This granularity leads to SCT channel occupancies below 1%.

Transition radiation tracker

While semiconductor trackers allow for a high granularity and a fast readout, they also come with disadvantages compared to conventional drift chamber trackers. They need to be cooled in order to suppress noise and they are more expensive. For these reasons the outermost component of the ATLAS inner tracking detector, the TRT, is made of drift tubes. It provides additional tracking information and supports the electron identification with the help of transition radiation (TR). The relatively low granularity of the TRT

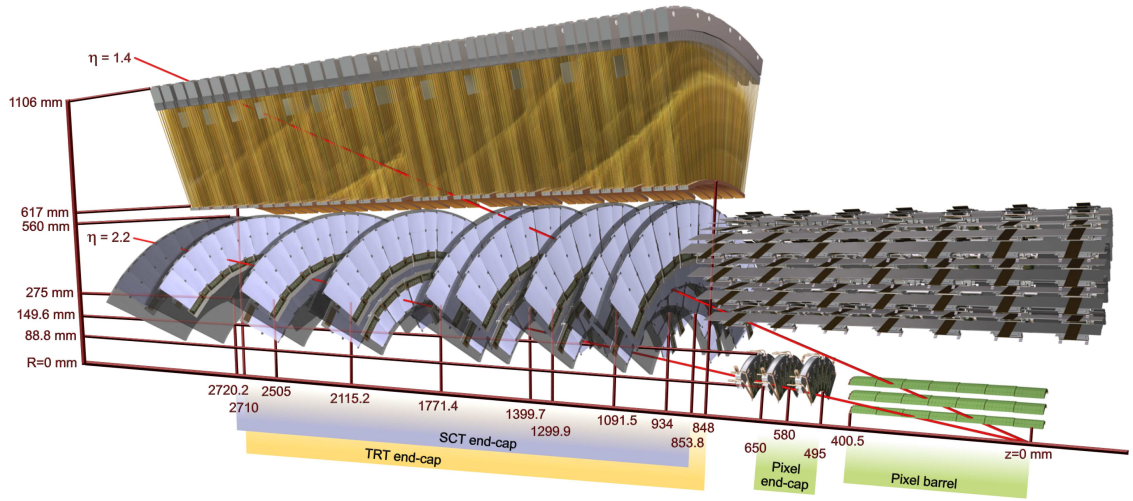


Figure 2.6.: ATLAS inner tracking detector, barrel and end-cap regions. Charged particles in the region $|\eta| < 2.5$ traverse several sensors of each of the three detector components. The red lines show two charged particle tracks of $p_{\mathrm{T}} = 10\,\mathrm{GeV}$ [141].

with an occupancy above 30% is compensated by a larger number of sensors with about 36 hits per track segment. The drift tubes, or straws, contained in the central part are aligned parallel to the beam axis. These straws are divided in half and read out at each end in order to reduce the occupancy. In the end-caps straws are installed radially. Each of the straws has a diameter of 4 mm and is made of a mixture of materials. In particular, polyamide is used as bare material, kapton is contained in the walls and carbon fibres serve mechanical stability. The anodes are 30 µm tungsten wires plated with gold. The employed gas is a mixture of 70% Xe serving TR absorption, 27% CO_2 ensuring high drift velocities and 3% O_2 increasing operation stability. The gas is circulating through the TRT while exit gas is being cleaned and recycled. The voltage of 1 530 V applied to each straw leads to a maximum drift time of about 50 ns. TR is produced by particles with high relativistic γ factors, in particular by electrons, traversing fibres and foils placed between the straws. TR signals are distinguished from charged particle ionisation signals by their higher amplitude using a dedicated readout threshold.

2.2.3. Calorimeters

In order to meet the various requirements of calorimetry, namely high energy and angular resolutions, efficient particle identification, hermetic coverage of solid angles and low rates for the punch-through of particles other than muons into the muon spectrometer, ATLAS

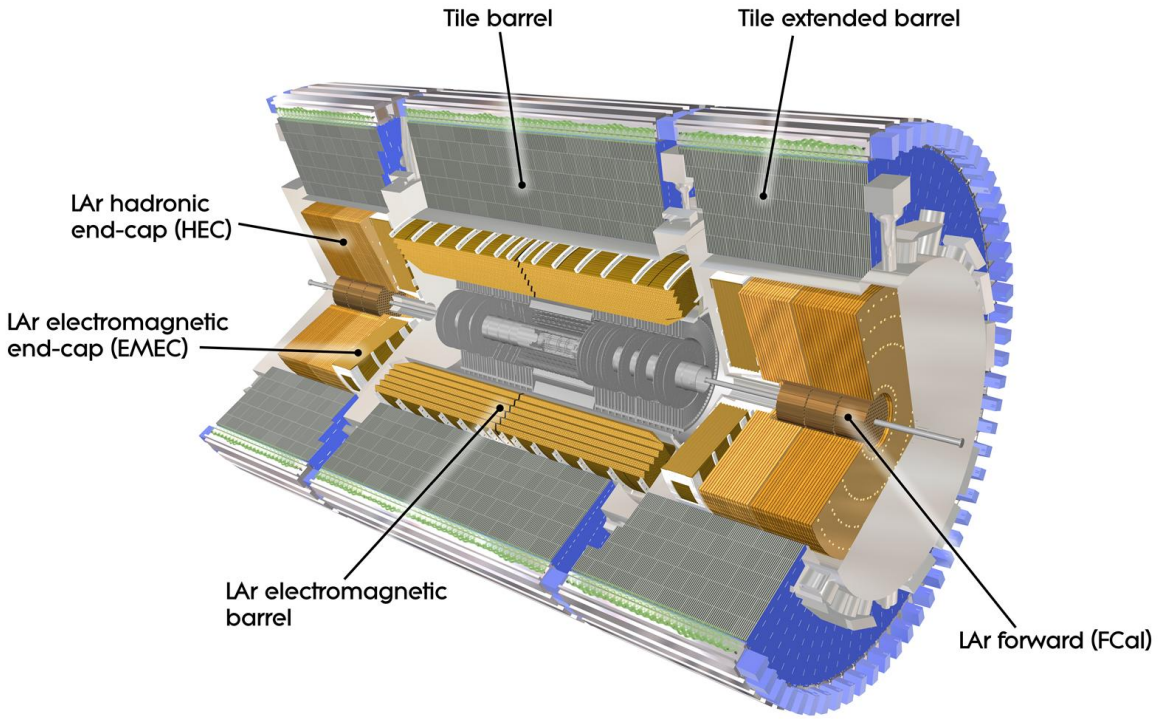


Figure 2.7.: ATLAS calorimeters. LAr denotes liquid argon. An electromagnetic calorimeter is followed by a hadronic calorimeter. Scatterings under small angles are covered by a dedicated forward calorimeter on each side. Therefore the IP is covered in almost all directions [139].

contains several types of calorimeters (Fig. 2.7). An electromagnetic (EM) calorimeter is installed downstream of the central solenoid magnet. It is followed by a hadronic barrel calorimeter. Their ranges are extended by electromagnetic and hadronic end-caps while there is a dedicated forward calorimeter on each side close to the beam pipe. In the following an overview of these calorimeters is presented.

Electromagnetic calorimeter

The EM calorimeter consists of a barrel covering a range of $|\eta| < 1.5$ and two end-caps covering a range of $1.4 < |\eta| < 3.2$. It has a sandwich structure with lead absorbers and liquid argon (LAr) chosen as a radiation hard active material. Each of the three calorimeter parts is contained in a dedicated cryostat. The end-caps are divided into two wheels with a gap in between at $|\eta| = 2.5$ which introduces a small non-sensitive region. The absorbers are shaped like an accordion, thus full coverage in the azimuthal direction

is achieved since no channelling cracks between absorbers are needed. Instead, copper electrodes, installed within the LAr layer between the absorbers, are read out at the front and at the rear of the calorimeter. A segmentation is realised by appropriate longitudinal divisions and readout ganging of the electrodes (Fig. 2.8). The short first layer with its high granularity in the η direction provides information about the direction of primary particles which is important e.g. for the measurement of prompt photon directions as well as their distinction from photon pairs originating from π_0 decays. Two more layers follow, while the second, central layer absorbs most of the energy. The segmentation of theses layers is chosen small enough to be able to distinguish between the thin electromagnetic and the broader hadronic showers. Overall the thickness of the EM calorimeter amounts to about 25 radiation lengths, denoted by X_0 in the following. The EM calorimeter is aided by a presampler installed in front of it, covering the range of $|\eta| < 1.8$. This short LAr layer equipped with electrodes allows one to correct for energy losses happening already upstream of the calorimeter. In fact, primary particles must traverse material of at least $2\mathrm{X}_0$ while there is a local peak of about $6\mathrm{X}_0$ at $|\eta| = 1.5$ before they reach the calorimeter.

Within the LAr active material, electrons set free by shower particles induce signals at the electrodes operated with voltages of about 2kV. As the corresponding pulse length of about 400 ns is much greater than the LHC bunch crossing gap, the pulses are shaped by the readout electronics, giving peak amplitudes after approximately 25 ns. The EM calorimeter provides an energy resolution of about

$$\frac{\sigma_E}{E}|_{\eta=0} \approx \sqrt{\left(\frac{9\%}{\sqrt{E[\mathrm{GeV}]}}\right)^2 + (0.5\%)^2} \tag{2.5}$$

for central primary particles. The first contribution is caused by statistical shower fluctuations while the second one is caused by imperfections of the calorimeter material. An additional noise term with a constant absolute uncertainty is neglected here. The calibration of the EM calorimeter is supported by prior electron test beam measurements.

Contrary to the tracking detectors, the calorimeter resolution improves with energy. The comparison of the Eq. 2.4 and 2.5 shows that beyond electron energies of about 30 GeV the EM calorimeter provides a better resolution than the inner tracking detectors.

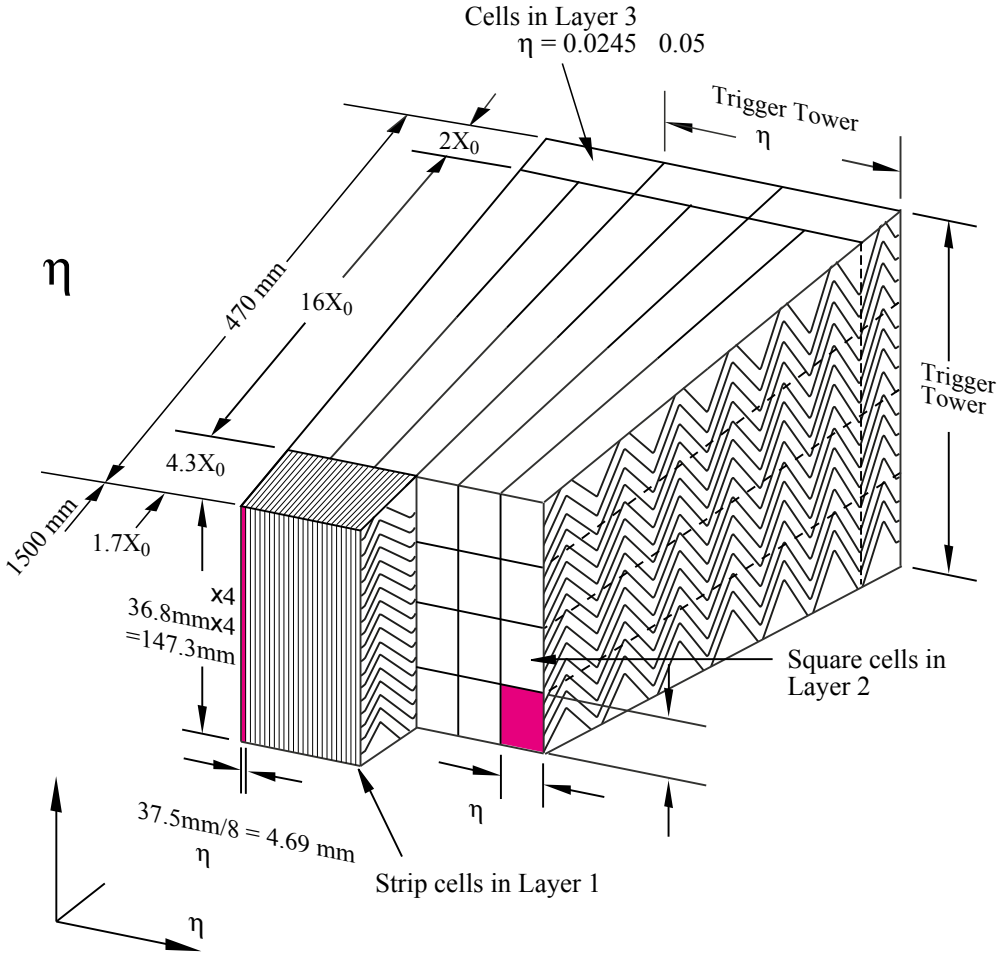


Figure 2.8.: A section of the ATLAS liquid argon electromagnetic calorimeter. X_0 denotes the radiation length. The accordion shape of the lead absorbers reduces the need for channelling. Cells are formed by the summation of readout channels [140].

Hadronic calorimeters

The hadronic calorimeters follow the EM calorimeter since hadronic showers are typically deeper than EM showers. Because of the increase of irradiation doses with pseudorapidity, different technologies are employed in different pseudorapidity regions. In the range $|\eta| < 1.7$ a tile calorimeter subdivided into a central barrel and two extended barrels is used. It has a sandwich structure made of steel absorbers and scintillating tiles used as the active medium. This choice was made in particular because it is less expensive than a continuation of the EM calorimeter technology. The radial depth of the tile calorimeter amounts to at least 10 interaction lengths which provides a sufficient containment of hadronic showers. It consists of modules in which several layers of tiles are placed (Fig. 2.9). The tiles are not facing the shower directions but are rather aligned

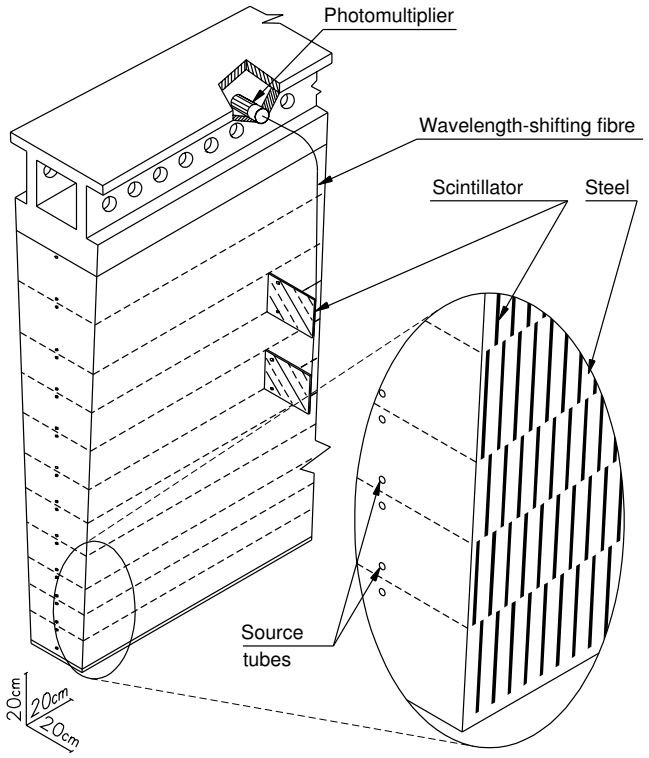


Figure 2.9.: ATLAS hadronic tile calorimeter module. Scintillating tiles are placed within steel absorbers. Their light is guided to wavelength-shifting fibres connected to photomultipliers installed on top of the module [140].

perpendicular to the beam axis which allows one to optimise hermeticity. The energy resolution is still sufficiently good because the velocities of low-energy charged particles emerging at the end of hadronic showers are distributed almost isotropically. The tiles are made of polystyrene which scintillates ultraviolet light. Wavelength-shifting fluoric fibres installed at the tile edges convert it into visible light which in turn is guided by the fibres to photomultiplier tubes (PMTs) installed at the top of the modules. The decay time of the wavelength-shifters amounts to about 6 ns. Fibres are grouped in order to form calorimeter cells which are facing the incident showers. The front-end electronics are installed next to the PMTs.

The range of $1.5 < |\eta| < 3.2$ is covered by hadronic end-cap calorimeters (HEC) which have a sandwich structure too. Copper is used as the absorber while LAr is used as the active material. The design of the HEC was driven by the aims of a simple mechanical construction, radiation hardness and relatively low costs [142]. The HEC shares a cryostat with the EM end-cap and forward calorimeters on each side. Both of the end-caps consist of two wheels which are longitudinally separated into two sections. The wheels are made

of wedge-shaped modules containing several layers of copper and LAr. Within the LAr gaps, electrodes are installed. Calorimeter cells pointing towards the IP are realised by a division of electrodes into pads and an appropriate summation of pad signals. Finally, readout electronics are mounted on the outer edges of the HEC.

Forward calorimeters

The region $3.1 < |\eta| < 4.9$ is covered by forward calorimeters (FCal) designed to deal with the high particle flux which occurs in this range. On each side there are three modules with an EM layer followed by two hadronic layers. In the sandwich structure of the FCal copper is chosen as the absorber of the first layer in order to optimize heat removal while in the second and third layer mainly tungsten is used as the absorber, which provides a sufficient containment of hadronic showers. For the purpose of shower detection, the absorbers contain holes which were drilled parallel to the beam axis. They contain concentric rods and tubes with LAr gaps in between used as the active material. These cylindric gaps are relatively small compared to the other LAr calorimeters of ATLAS as this setup provides faster signals and avoids large space charges caused by ions. The rods serve as anodes while the tubes and metal matrices are grounded. In order to keep electronics away from the region of largest radiation damage which is close to the back of the first layer, signals are read out at the front of the first layer and at the back sides of the second and third layers.

2.2.4. Muon spectrometer

Muons are the only particles, aside from the weakly interacting neutrinos, which can travel long distances in dense material. Hence their detection downstream of the calorimeters is an important technique of event reconstruction. For this reason and due to the occurrence of muons in many interesting electroweak scattering processes, special effort has been made to provide triggers which indicate the presence of muons and to precisely measure muon momenta. The ATLAS muon spectrometer (MS) follows the calorimeters and covers the range $|\eta| < 2.7$, while muon triggers are restricted to $|\eta| < 2.4$. It combines several techniques of muon detection and is supported by the toroid magnets already described in Sec. 2.2.1 (Fig. 2.11).

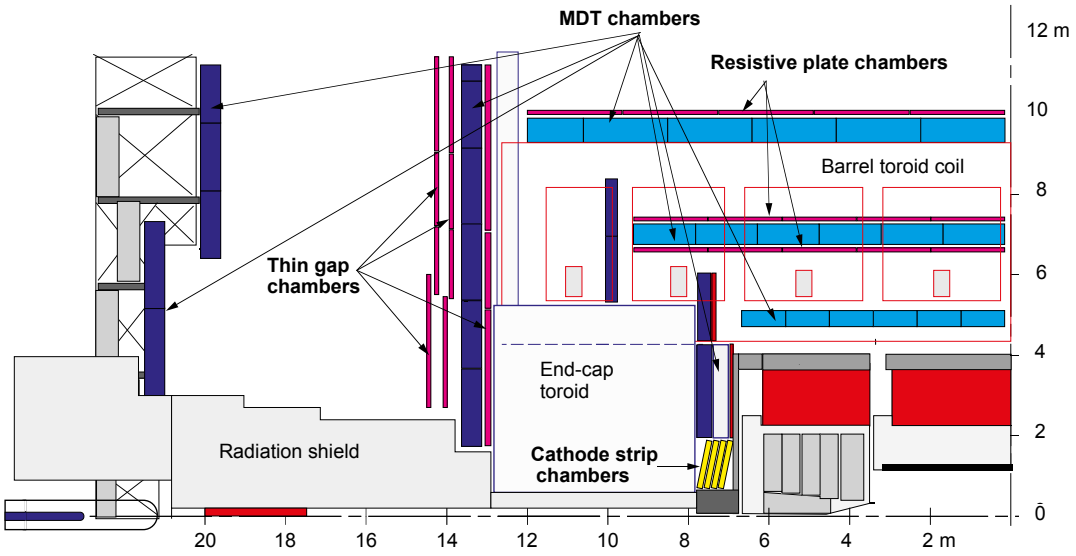


Figure 2.10.: Side view of the ATLAS muon spectrometer. Monitored drift tubes (MDT) and cathode strip chambers are used for muon precision tracking. Resistive plate chambers and thin gap chambers complete the tracking information and provide fast muon trigger signals [144].

Monitored drift tubes (MDTs) measure muon track coordinates in the bending direction perpendicular to the toroidal magnetic field. Particularly radiation hard cathode-strip chambers (CSCs) are used for precision tracking in the forward direction. Additionally, resistive plate chambers (RPCs) and thin gap chambers (TGCs) serve the purposes of muon triggering, the assignment of bunch crossings to muons and the measurement of azimuthal coordinates of muon tracks.

In the central barrel the MS consists of three cylindrical layers of MDTs surrounded by RPCs. In the end-caps there are several wheels installed at different distances from the IP. Again, MDTs are used except for the innermost wheel in the forward direction $2.0 < |\eta| < 2.7$ where CSCs take over. In addition, there are MDTs parallel to the beam axis located at the outer edge of the end-cap toroids. TGCs are used instead of RPCs in the end-caps. While the geometric coverage of the trigger chambers amounts to about 99% in the end-cap regions, it is reduced to about 80% in the barrel region due to services necessary for the inner tracking detector and calorimeters, support structures and two small elevator shafts in the lower part of the MS [143].

As the momentum resolution requirements surpass the mechanical accuracy, the MDTs as well as the CSCs are permanently monitored by an optical alignment system. The position of chambers relative to each other as well as deformations occurring within chambers are detected as deviations from straight lines. However, the optical alignment

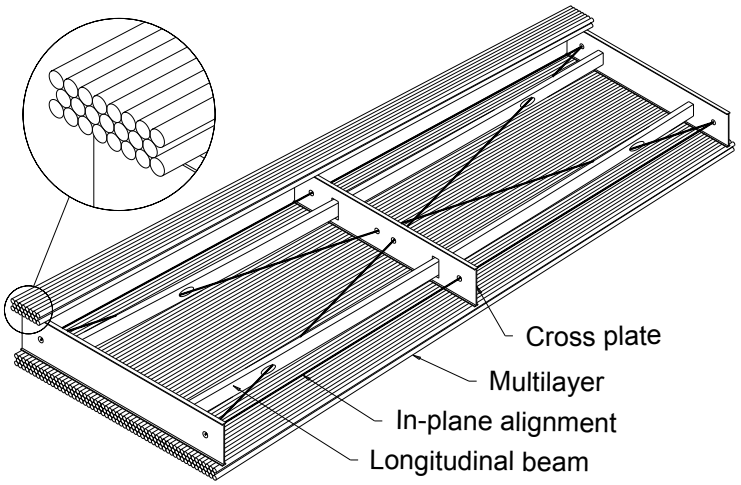


Figure 2.11.: Example of an ATLAS MDT module. Two multilayers of drift tubes provide precise tracking information in the magnetic field bending plane. The alignment is monitored permanently by optical means [144].

system is not capable of determining the position of MS chambers relative to the calorimeters and to the inner tracking detector. For this purpose it is aided by an alignment based on measured tracks.

Monitored drift tubes

The high precision tracking MDTs consist of two multilayers of Al drift tubes of 3 cm diameter, placed on aluminium spacers (Fig. 2.10). The forms of the MDTs differ depending on their position. While the mechanical construction is relatively simple, it allows only for small tolerances. Gold-plated tungsten-rhenium wires of 50 µm diameter serve as anodes and a voltage of 3 kV is applied. A gas mixture of 93% Ar and 7% CO_2 with a pressure of 3 bar is used, which generates basically no deposits in operation. The maximal drift time amounts to 700 ns. Tube signals are passed on to amplifiers, shapers and discriminators, and are subsequently fed into time-to-digital converters.

Cathode-strip chambers

For the purpose of precision tracking in the inner forward region of the MS, which is affected most strongly by irradiation, CSCs are used. They are multi-wire proportional chambers of 5 mm thickness arranged in four layers on each side. The cathodes surrounding the anode wires are divided into strips and serve the signal readout. In order

to obtain a two-dimensional resolution, opposing cathode planes are segmented in two directions perpendicular to each other. Operated with a different Ar/CO_2 gas mixture and using charge interpolation between cathode strips, the CSCs provide fast and precise tracking information.

Resistive plate chambers

The RPCs contain layers of two parallel electrode planes while the gap between the electrodes is filled with gas. As no wires are used, the RPCs are simple to manufacture and provide a good time resolution. Unlike wire chambers, in which primary ionisation clusters arrive in sequence at the multiplication region near the wire, the uniform electric fields of the RPCs lead to simultaneous avalanches giving a single signal with low variance in time. These signals are read out with the help of metallic pick-up strips placed on the outer faces of the electrodes. These sets of strips are aligned orthogonal to each other in order to allow for measurements in the bending and azimuthal plane.

Thin gap chambers

Multi-wire proportional chambers provide muon trigger input from the forward regions which are strongly affected by irradiation. These TPCs are made of high resistivity cathode planes with anode wire planes placed in between. The wires are grouped depending on the pseudorapidity region and provide the coordinate measurement in the bending plane. Some of the copper claddings at the back of the cathode planes are segmented into strips. They are read out too and provide the azimuthal coordinate. Due to strong electric fields close to the wires and the small distance between wires, the TGCs achieve a good time resolution.

2.2.5. Electronics, trigger and data acquisition

The LHC bunch crossing rate of up to 40 MHz is a challenge for the acquisition of collision data. There is not enough storage capacity available for the experimental collaborations to store every collision event. In addition, the complex detectors cannot be read out completely at this high pace. For these reasons an in situ selection of events must be performed. This technique, called triggering, does not compromise the goals of physics

analyses as the most interesting processes occur only at a rate much lower than 40 MHz. On the other hand, abundant processes which are of interest too, e.g. the production of jets in strong interactions, can be studied with the help of triggers for which the rate is artificially suppressed by a defined factor, called the pre-scale.

The ATLAS trigger system is divided into three levels as illustrated in Fig. 2.12. Going from one level to another, the event rate is steadily reduced whereas the applied reconstruction algorithms become more accurate. The Level-1 trigger (LVL1) is implemented as a system of dedicated electronics, while the Higher-Level trigger, consisting of the Level-2 and the final Event Filter (EF), is implemented on PC farms. The trigger system is configured according to a menu of various signatures of interest. If any of these trigger items fires, the event is selected.[1]

Electronics

The first stage of data readout is given by the front-end electronics located inside the ATLAS detector itself. Analogue-to-digital signal conversion is mostly done at this point. Pipeline buffers save the signals of all bunch crossings during the LVL1 decision latency. Clearly, for this mechanism to work the buffers and the LVL1 trigger must by synchronized. For this purpose, timing, trigger and control links (TTC) provide the LHC beam clock which is synchronised with the LHC machine. The phase of the clock with respect to the bunch crossings is determined with the help of beam position monitors located 175 m away from the ATLAS IP on each side. In addition to the beam clock, the TTC system also sends the LVL1 accept signal (L1A) back to the front-end electronics if the LVL1 has fired and it provides different control signals like test pulses or reset commands. The back-end electronics, including the LVL1 trigger, are located in counting rooms close to the detector cavern while the data acquisition (DAQ) room is located in a surface building (Fig. 2.3).

Level-1 trigger

The first trigger level consists of a central trigger processor which makes use of trigger information delivered by a calorimeter trigger system (L1Calo) and a muon trigger

[1] Meanwhile the Level-2 and the Event Filter have been combined into a single stage. However, the discussion presented here refers to the trigger architecture chosen in 2012 and earlier, which is relevant to the measurement presented in this thesis.

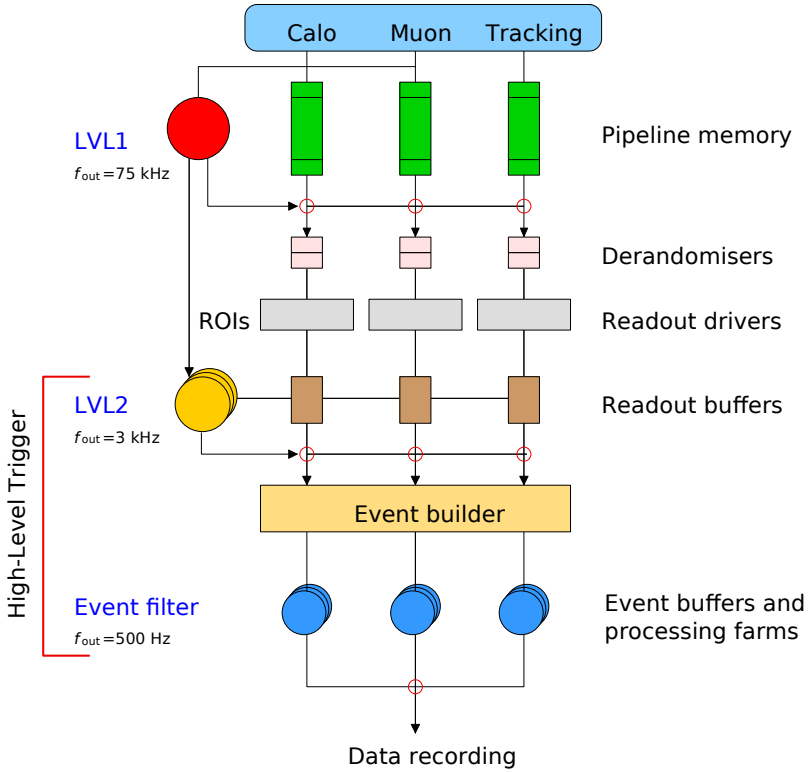


Figure 2.12.: Scheme of the ATLAS trigger system. Three trigger levels gradually reduce the event rate while the accuracy of the algorithms used to identify interesting events increases. Signals of selected events are read out of pipelined buffers integrated into the front-end electronics, stored in readout buffers for further trigger access, put into a common memory before the arrival of the final trigger decision and eventually sent to permanent storage at the CERN computing center [3].

system (L1MU). While L1MU uses signals from the RPCs and TGCs, L1Calo refers to all calorimeters. Calorimeter cells are combined to form trigger towers of reduced granularity. At LVL1, the presence of high-p_{T} objects as well as large missing and total transverse momentum is investigated. Several p_{T} thresholds are involved and different objects can be combined. The input rate of 40 MHz is reduced to an L1A rate of about 75 kHz while decisions require a latency of about 2 µs. In case of an L1A decision, the signals saved in the pipeline buffers at the front-end electronics are multiplexed and sent to readout buffers (ROBs) via readout drivers (RODs). Intermediate buffers, called derandomisers, average out the random L1A rate in order to match the ROD input bandwidth. An L1A signal causes a short readout dead time of about 150 ns resulting in a small loss of data. This is taken into account when the integrated luminosity of the data set is determined later on. In addition to the trigger decision, regions of interest

(RoIs) are identified by LVL1. They correspond to momenta of objects which triggered the L1A decision and are sent to LVL2.

Higher-Level trigger

The LVL2 selectively analyses event data stored in the ROBs. As this selection is based on the RoI information, typically only a small amount of data on the order of 1% needs to be read out. An improved event selection is achieved by several means. For instance, higher granularity data is used, precision tracking detectors come into play, further isolation requirements are applied and electron tracks can be matched to calorimeter clusters. Overall, the event rate is reduced to about 3 kHz with a decision latency of about 40 ms.

In case of a positive LVL2 decision, the event data stored in the ROBs is moved forward and combined in a single memory for the first time. This process is called event building. As a last triggering step, the EF is applied to these events. It benefits from more refined algorithms and the latest calibration and alignment information. Also with a decision latency of about 4 s more processing time is available compared to LVL2. The output rate of events sent to permanent storage is on the order of a few 100 Hz. In addition to the event selection, the EF assigns the selected events to streams (*EGamma*, *Muon*, *JetEtMiss* among others) which simplifies the following physics analysis.

Data storage and distribution

All events selected for permanent storage are sent to the CERN computing center for offline reconstruction. In order to store and distribute the data of LHC experiments, the Worldwide LHC Computing Grid (WLCG) was created. Its hierarchy is made of four so-called Tiers with Tier 0 being the CERN computing center whereas the other Tiers are spread around the world. Thirteen large computing centers serve as the Tier 1 level while about 160 further sites serve as Tier 2. The end-user benefits from a fast and reliable data access. Furthermore, the WLCG provides considerable computing power [145].

2.3. LHC and ATLAS performance in 2012

After first considerations started in the year 1984, the decision to build the LHC was made by the member states of CERN in 1994. The design, construction and commissioning of

the collider and its detectors was completed in September 2008 when the first proton beams were steered around the LHC. However, only a few days later an accident caused a one-year break of the machine operation. A faulty magnet connection had caused resistivity which led to a production of heat. Helium was suddenly released and caused serious damage to magnets. After the necessary repairs and the installation of improved protection systems, the LHC realized first pp collisions in November 2009. This was the start of its first long operation period which lasted until early 2013. Then a first long shutdown began in order to prepare the collider for higher collision energies.

In this thesis pp collision data recorded by ATLAS in the year 2012 is analysed. The LHC was operated with a bunch gap of 50 ns, thereby reaching instantaneous luminosities of up to $7.73 \cdot 10^{33}$ cm^{-2}s^{-1}. The luminosities were determined with the help of van der Meer scans [146]. In these dedicated runs of operation, the colliding beams are shifted relative to each other in order to determine their density profile. Together with beam current measurements, this information allows one to determine the absolute luminosity $\mathcal{L}$ (Eq. 2.2). In addition, luminosity detectors measure the number of inelastic pp interactions per bunch crossing within their acceptance, μ_{pp}^{vis}, which is related to the inelastic pp interaction rate $\dot{N}_{pp}^{\text{vis}}$. From $\mathcal{L}$ and $\dot{N}_{pp}^{\text{vis}}$, the total inelastic pp cross section σ_{pp}^{vis} can be derived. This is the calibration constant necessary to determine the absolute luminosity from the observed interaction rate during the nominal data taking, $\mathcal{L} = \dot{N}_{pp}^{\text{vis}} / \sigma_{pp}^{\text{vis}}$. In ATLAS this technique is applied for online monitoring as well as for offline physics analyses.

In 2012 the LHC and ATLAS usually operated smoothly. All ATLAS subdetectors described above were working with an operational fraction of channels close to 100%. Typical trigger rates were on the order of 50 kHz, 4 kHz and 500 Hz for LVL1, LVL2 and the EF, respectively. The EF output rate was limited by the computing power at Tier 0, where the offline reconstruction of all selected events had to be performed. Some of the selected events were stored for later reconstruction in order to allow for a higher EF output rate.

ATLAS recorded most of the collision events which were delivered to it by the LHC. The resulting data set of 2012 amounts to an integrated luminosity of about 20 fb^{-1} (Fig. 2.13). Concerning colliders at the energy frontier, this is a data set of unprecedented volume and energy. It allows for numerous new insights into high energy physics, one of them being the single top-quark measurement presented in this thesis.

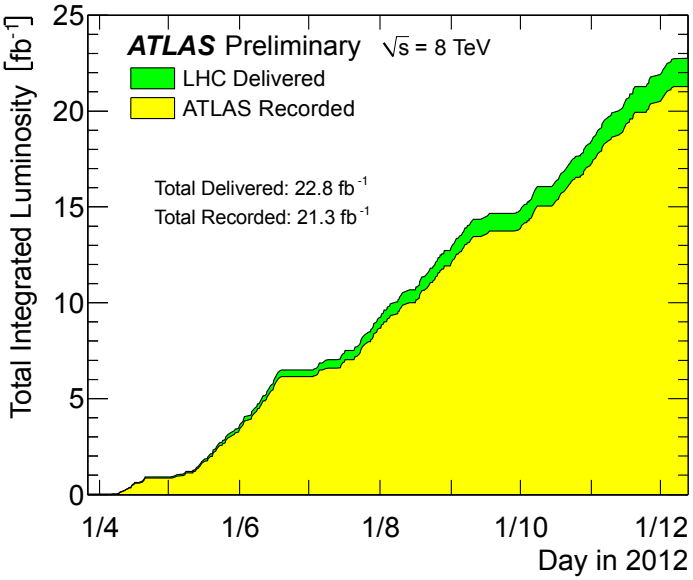


Figure 2.13.: ATLAS integrated luminosity in 2012. For several months the LHC delivered pp collisions and ATLAS recorded most of them. The data set used in this thesis adds up to $20.3\,\mathrm{fb}^{-1}$. Due to quality selection cuts it is a subset of the total amount of recorded data which adds up to $21.3\,\mathrm{fb}^{-1}$ [139].

2.4. ATLAS event simulation

The collision events at the LHC result in complex detector responses. Regarding physics analyses, there are two ways to cope with this condition. Either an analysis is designed in a way which makes sure that only a few aspects of the detector response matter, or many details of this response need to be understood by means of a simulation. An example of an analysis which has only a small dependence on detector simulations is the search for the standard model Higgs boson in the di-photon decay channel $H \to \gamma\gamma$, where simply a peak in the di-photon invariant mass spectrum is identified [1]. However, the modelling of various processes by means of detector simulations is by far the more frequent approach, in particular in the context of analyses of top-quarks with their relatively complicated decay signatures.

The ATLAS event simulation proceeds in three major steps [147]. First, pp collisions with final states of hadrons and leptons are simulated as described in Sec. 1.4. This is followed by a simulation of the response of the various detector sensors to the primary particles. Finally, the response of the detector electronics to these sensor hits is simulated. This event simulation chain is embedded in the ATLAS software and large-scale

productions of MC simulation samples are performed on the WLCG (Sec. 2.2.5). The resulting simulated events can be analysed in the same way as measured data events.

Pile-up and reweighting

Apart from the hard scattering event, several other effects need to be taken into account. The corresponding hits in the sensors are added to those of the hard scattering process. There are interactions of beam particles with residual gas contained in the beam pipe and with material of accelerator structures. There is radioactivity caused by neutrons and kaons as well as a photon background in the ATLAS cavern. Interactions from neighbouring bunch crossings matter too, as some signal integration times extend over many bunch crossings. Most notably, multiple *pp* interactions occur within one bunch crossing. These are soft interactions, but they need to be taken into account. While "pile-up" stands for the sum of these effects, it will only denote the latter effect in the following.

For each of the simulated hard scattering events, a varying number of pile-up interactions is added. In order to match the distribution of the number of pile-up events to measured data, a reweighting is applied. Figure 2.14 shows the distribution of the mean number of interactions per crossing, denoted by μ, at the ATLAS IP in 2012, weighted to the luminosity recorded with the corresponding μ values. The aim of the pile-up reweighting is to match the simulated distribution of interactions per crossing to this data distribution. The general idea of reweighting is to apply weights w_i to simulated events in order to fit a simulated distribution to data. Let N_i^{data} and N_i^{MC} be the number of real and simulated events in bin i, respectively,

$$w_i \coloneqq \frac{N_i^{\text{data}}}{N_i^{\text{MC}}}, \quad N_i^{\text{MC, w}} \coloneqq w_i \cdot N_i^{\text{MC}} \;\Rightarrow\; N_i^{\text{MC, w}} = N_i^{\text{data}} \; . \tag{2.6}$$

The reweighted distribution $\{N_i^{\text{MC, w}}\}$ matches the data. Concerning the pile-up reweighting, weights for events simulated with a certain μ value are given by the ratio of the relative amount of integrated luminosity corresponding to this μ value, L_μ/L, and the relative fraction of events generated with it, N_μ/N: $w_\mu = (L_\mu/L) \,/\, (N_\mu/N)$. More precisely, this procedure is carried out after a splitting of the data set into several run periods, each of which involves approximately the same instantaneous luminosities and hence similar pile-up conditions.

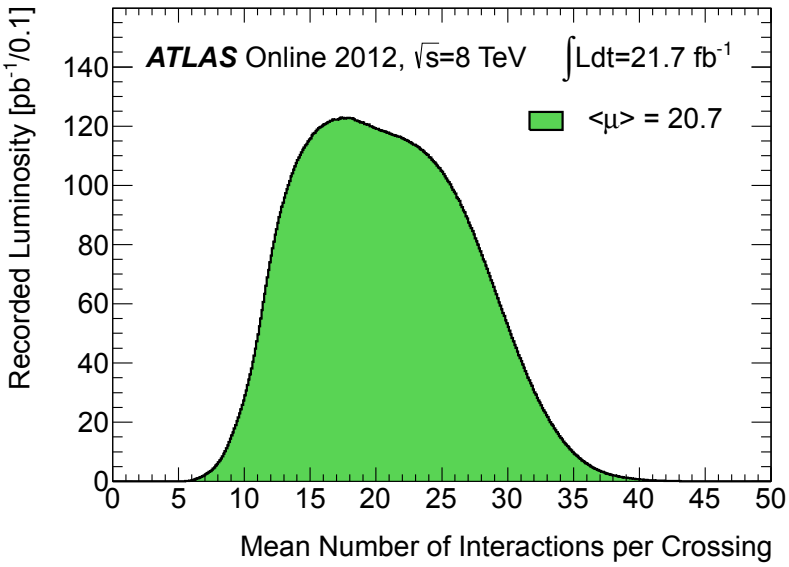


Figure 2.14.: Number of interactions per bunch crossing in 2012. While the number of pp interactions in a bunch crossing follows a Poisson distribution, the mean number of interactions per bunch crossing μ is shown here. The μ distribution is weighted with the integrated luminosity which was recorded with each particular μ value [139].

Simulation of sensor responses

Primary particles are propagated through a simulation of the ATLAS detector. The decays of long-lived particles are simulated here as well. The ATLAS detector information relevant to simulations is stored in databases. There is a construction database which provides the detector geometry and material information, a magnetic field map, and a conditions database providing e.g. calibrations, alignments and dead channel information. The interactions of primary particles with the detector are simulated using the GEANT4 toolkit [148]. It allows one to simulate various interactions in different materials for a wide range of particle energies. This is the most time-consuming simulation step. Several minutes are needed to simulate a single event. Alternatively, the fast simulation program ATLFAST-II can be employed. It saves time by using simple parametrisations of calorimeter responses. However, GEANT4 is usually preferred.

Digitization

The simulation of sensor responses is followed by a simulation of the corresponding behaviour of the detector electronics. The generated sensor hits are converted to voltages

and currents. Features of the electronics like cross-talk or noise are taken into account. This digitization is tuned to reproduce data received from laboratory tests, test beams and cosmic ray data taking. The LVL1 trigger behaviour with responses of the trigger items is simulated as well. However, no events are discarded in simulations.

2.5. Physics object reconstruction

While in the previous sections the gathering of raw data representing hits in the various detector sensors was discussed, the reconstruction of physics objects like electrons, jets etc. from these hits is the next and last step which needs to be taken before the physics of the primary interactions can be studied. The top-quark events analysed in this thesis involve several objects, namely electrons, muons, jets induced by b-quarks and missing transverse momentum caused by neutrinos. In this section an overview of the reconstruction of these objects as performed in 2012 is given, starting with primary vertices as their common origin.

2.5.1. Primary vertices

The pp collisions delivered to ATLAS in 2012 involve multiple pp interactions as discussed in Sec. 2.4. Hard interactions are indicated by a vertex at which several high-p_{T} tracks intersect. Furthermore, the identification of b-jets, which is important for top-quark analyses expecting SM top-quark decays, involves the reconstruction of secondary vertices which need to be assigned to a primary vertex (PV). For these reasons, PVs must be reconstructed.

First, tracks are reconstructed in the inner detector with an "inside-out" algorithm. Silicon tracker hits are used as seeds and more hits with increasing distances from the IP are added to a track. Outcomes are accepted as tracks if they contain a minimum number of hits in the silicon detectors, while no lack of a hit in an active pixel detector module traversed by the track is allowed [149]. PVs are reconstructed from these tracks with the help of an iterative algorithm. The longitudinal components of tracks at the beam line are used in order to find vertex seeds. Following this, χ^2 vertex fits are performed using tracks found in the vicinity of the seeds. Incompatible tracks are removed from the corresponding vertex and serve as new seeds. This procedure is repeated until no vertices are left [150].

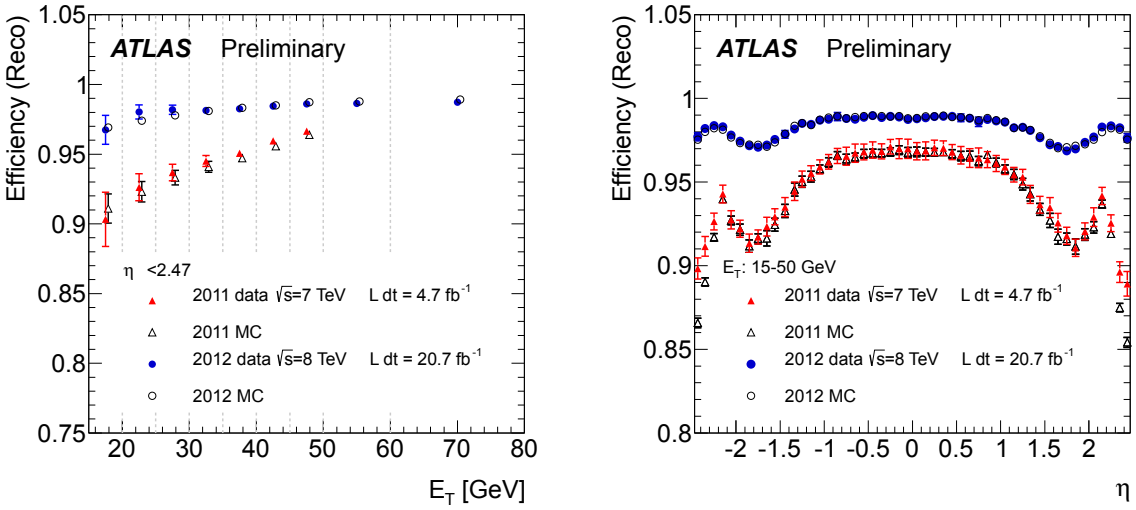
2.5.2. Electrons

Several steps are taken in order to detect electrons. First of all, energy clusters deposited in the EM calorimeter are matched to tracks. This step is called "reconstruction" in the following. Electron candidate energies are calibrated and further cuts are applied in order to purify the selection. Also possible overlaps between electrons and other high-p_{T} objects must be resolved. In addition, the performance of the overall electron reconstruction in simulations must be matched to the performance given in measured data. The following section presents an overview of these procedures for electrons within the inner tracking detector acceptance. In addition, the electron trigger relevant to the analysis presented in this thesis is discussed.

Reconstruction

The mere reconstruction of electrons starts with a search for energy clusters in the EM calorimeter using a "sliding-window" algorithm [151]. Calorimeter cells are summed up to towers of reduced granularity while the tower energy is given by the sum of all corresponding cells. Connected sets of towers, called "windows", are scanned for local maxima of E_{T} while the total E_{T} needs to pass a certain threshold. Then, inner detector tracks pointing to the barycentre of such seed clusters within a distance of $\Delta R = 0.3$ are being searched for.

The reconstruction of tracks is performed with an "inside-out" algorithm as mentioned in Sec. 2.5.1. Sensor hits picked with regard to a possible track serve as input to a track fit. After the fit, a loose matching between tracks and seed clusters is required. Matched tracks are refitted using an optimized algorithm, the Gaussian Sum Filter [152], which is a generalization of the Kalman Filter algorithm. The track-cluster matching is repeated using these improved track fits and tighter matching requirements. If more than one track is matched to a cluster, a primary track is chosen on the basis of quality characteristics like the presence of hits in the pixel detector or the ΔR distance between the track position at the EM calorimeter and the cluster. The cluster formation is repeated and the energy calibration is applied. For the final electron candidates, the directions are given by the track except for TRT-only tracks, while the energy is given by the cluster energy as the EM calorimeter surpasses the inner tracking detector in terms of energy resolution for high-E_{T} electrons (see Sec. 2.2.3). Examples of measured reconstruction



(a) **(b)**

Figure 2.15.: ATLAS electron reconstruction efficiency. In (a) the measured efficiencies are shown as a function of E_{T}, integrated over pseudorapidities, while (b) refers to different pseudorapidities for $15\,\mathrm{GeV} < E_{\mathrm{T}} < 50\,\mathrm{GeV}$. The measurements were performed with tag-and-probe methods as described at the end of this section. Efficiencies concerning the 2012 data set (circles) are improved compared to 2011 (triangles) and are close to one. The dashed lines in (a) indicate the binning which is applied for the efficiency calculation [151].

efficiencies are presented in Fig. 2.15. The applied measurement strategy is discussed at the end of this section.

Energy calibration

While test beam data allows one to convert signals into corresponding energy deposits, it is still necessary to reconstruct original electron energies at the IP. The energies of identified clusters must be corrected for energy loss upstream of the EM calorimeter and for energy leakage which reduces the deposited cluster energy [153]. For this purpose, EM cluster properties together with inner tracking information is mapped to a calibrated cluster energy according to results obtained from simulations. Multivariate techniques

are used in order to handle various input variables of the calibration studies like the overall measured energy, the energy measured in the presampler or the shower depth. The amount of material upstream of the EM calorimeter is estimated with the help of measured EM shower shapes. The resulting calibration factors depend on η and p_{T}.

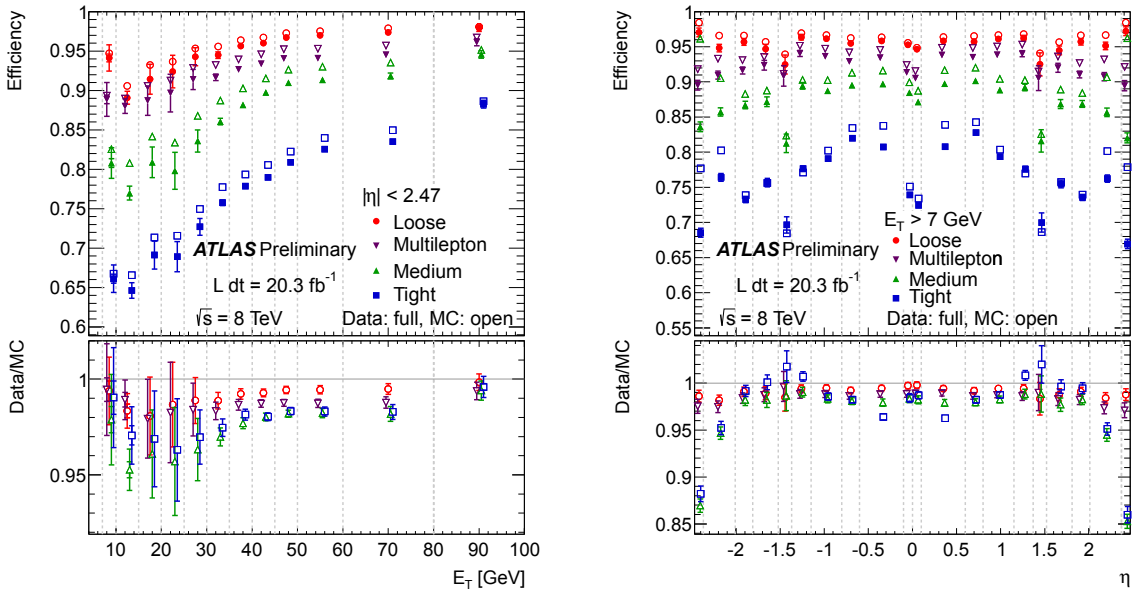
Furthermore, the energy scale of electrons is set with the help of measured $Z \to ee$ events as the mass of the Z boson is known precisely. The electron energy resolution is determined in this context as well. The calibration results in uncertainties for central electrons with $E_{\mathrm{T}} \approx 40\,\mathrm{GeV}$ of about 0.04% concerning the energy scale and less than 10% concerning the energy resolution.

Identification

In the set of reconstructed electron candidates, real, prompt electrons must be identified and distinguished from backgrounds like hadronic jets or secondary electrons created in photon conversions. For this purpose, a menu of discriminating variables with three classes of different prompt electron purity and efficiency is employed. There are loose, medium and tight selections, where tighter selections are subsets of looser selections. Several variables concerning e.g. EM calorimeter shower shapes, track properties or matching degrees between tracks and energy clusters are used. Tighter selections are characterized by a larger number of variables and tighter cuts. For instance, the amount of hadronic leakage is considered in all selections, pixel detector B-layer hits are required only in medium and tight selections, and the ratio E/p of the cluster energy E over the track momentum p is only used in the tight selection. Figure 2.16 presents measured identification efficiencies.

In order to improve the separation of electrons from hadronic jets, additional cuts are applied which require the electron candidate to be isolated from other high-E_{T} or high-p_{T} objects. The variable $E_{\mathrm{T}}^{\mathrm{cone}\,\Delta R}$ represents the sum of calorimeter cell transverse energies in a cone of a size of ΔR around the cluster barycentre, excluding the centre of the cluster. Here, a correction for additional pile-up contributions is applied, depending on the number of reconstructed primary vertices. There is a similar variable concerning the isolation of the candidate track, $p_{\mathrm{T}}^{\mathrm{cone}\,\Delta R}$. Considered tracks must originate from the assigned primary vertex, have $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ and meet minimum quality criteria.

In order to resolve overlaps between electrons and other high-p_{T} objects, only events in which no electron candidate shares a track with a muon are considered, at least as far



(a) **(b)**

Figure 2.16.: ATLAS electron identification efficiency. In (a) and (b) they are shown as a function of E_{T} and η, respectively. Isolation requirements are not included. Tag-and-probe methods were employed to perform the measurements, as explained at the end of this section. Different selections are compared to each other. The most pure selection, called "tight", which is most relevant to offline analyses, provides the smallest efficiencies. They are considerably lower than the reconstruction efficiencies presented in Fig. 2.15. The lower panels compare the measured efficiencies with their counterparts obtained in simulations before a reweighting has been applied. The dashed lines indicate the binning used for the efficiency calculation [151].

as top-quark analyses are concerned. Furthermore, electron candidates are discarded if they are close to a jet ($\Delta R < 0.4$).

Electron Trigger

The main single electron trigger used by ATLAS in 2012 is called `EF_e24vhi_medium1`, which has the lowest E_{T} threshold among all unprescaled single electron triggers used in that year. It is seeded on LVL1 by the `L1_EM18VH` item. Here `V` stands for "varied threshold" which means that the E_{T} cut of about 18 GeV is varied depending on the

pseudorapidity of the object in question, while `H` stands for "hadronic core isolation" which means that only a limited hadronic energy leakage of the EM cluster is accepted. The `i` in the EF trigger item reflects a tracking isolation requirement. The `24` in the EF item denotes the E_{T} beyond which the trigger efficiency reaches a plateau and is close to one, while `medium1` represents the identification cuts applied similar to the offline identification discussed above. Inefficiencies of the `EF_e24vhi_medium1` trigger occuring at high transverse energies, caused e.g. by the hadronic leakage cut, are compensated by the use of an additional trigger with a higher p_{T} threshold, called `EF_e60_medium1`.

Efficiency measurements and reweighting of simulated events

In order to properly model collision events, the overall electron reconstruction efficiencies in simulations must be matched to the actual efficiencies obtained from collision data. Therefore a reweighting similar to the pile-up reweighting introduced in Sec. 2.4 is applied for all simulated events concerning trigger, reconstruction and identification efficiencies. Accordingly, the overall weight is given by the product of these contributions,

$$\varepsilon_{\mathrm{total}} = \varepsilon_{\mathrm{trigger}} \cdot \varepsilon_{\mathrm{reco}} \cdot \varepsilon_{\mathrm{id}} \ , \tag{2.7}$$

$$\begin{aligned} \varepsilon_{\mathrm{total}}^{\mathrm{MC,\,w}} &= w_{\mathrm{trigger}} \cdot w_{\mathrm{reco}} \cdot w_{\mathrm{id}} \cdot \varepsilon_{\mathrm{total}}^{\mathrm{MC}} \\ &= \frac{\varepsilon_{\mathrm{trigger}}^{\mathrm{data}}}{\varepsilon_{\mathrm{trigger}}^{\mathrm{MC}}} \cdot \frac{\varepsilon_{\mathrm{reco}}^{\mathrm{data}}}{\varepsilon_{\mathrm{reco}}^{\mathrm{MC}}} \cdot \frac{\varepsilon_{\mathrm{id}}^{\mathrm{data}}}{\varepsilon_{\mathrm{id}}^{\mathrm{MC}}} \cdot \varepsilon_{\mathrm{total}}^{\mathrm{MC}} \\ &= \varepsilon_{\mathrm{total}}^{\mathrm{data}} \ . \end{aligned} \tag{2.8}$$

The scale factors (SFs) $\varepsilon^{\mathrm{data}}/\varepsilon^{\mathrm{MC}}$ are measured on the basis of clean and nearly unbiased electron samples obtained using the tag-and-probe method. In this context events containing electron pairs are selected, in particular $Z \to ee$ events. Strict quality conditions must be met by one of the electrons, called "tag", while the other electron, called "probe", is used for efficiency measurements. In order to avoid a bias of the probe selection, both permutations of the electron pair can be considered. Possible biases introduced by probe requirements or by the event selection are present in data as well as in simulations and are expected to cancel in the SFs to a large extent. The efficiencies are given by the fraction of probe electrons passing the corresponding selection criteria. As part of this procedure, background contributions to the probe selection like hadronic jets faking electron signatures or secondary electrons, created e.g. in photon

conversions, must be estimated. For this purpose, discriminating variables like the di-electron invariant mass or the calorimeter isolation $E_{\mathrm{T}}^{\mathrm{cone}\,0.3}$ of probe electrons are used. Comparisons of the different methods allow one to estimate systematic uncertainties of the efficiency measurements. The efficiencies presented in this section show results of these tag-and-probe measurements (Fig. 2.15, 2.16).

2.5.3. Muons

The muon reconstruction chain consists of several steps similar to the electron reconstruction discussed above. However, there is a relatively broad spectrum of muon reconstruction algorithms applied in ATLAS. The following discussion focusses on aspects which are relevant to the muon reconstruction applied in the context of the measurement presented in this thesis. Emphasis is put on the reconstruction of muons with $p_{\mathrm{T}} < 120\,\mathrm{GeV}$ within the inner tracking detector (ID) acceptance of $|\eta| < 2.5$ [154]. The relevant muon trigger is discussed as well.

Reconstruction and identification

Several "types" of muons are distinguished by the ATLAS muon reconstruction. They differ by the detector components which are involved in their reconstruction. The most common type is defined by combined muons for which tracks of the inner tracking detector (ID) and the MS are combined. Other types can serve for the purpose of extending the acceptance beyond the reach of this combined muon reconstruction at the cost of lower purity. Two reconstruction algorithms are available, namely the STACO and Muid algorithms. Concerning combined muons, in the STACO case track parameters are combined taking into account the covariance matrices of the ID and MS tracks, while in the Muid case a global refit of the muon track including ID and MS hits is performed. The latter approach is chosen for the measurement presented in this thesis.

Several quality levels are defined for Muid muons. For tight muons, the combined fit of the ID and MS tracks must be successful. Medium muons only involve MS tracks and loose muons are given by any muon type which involves an ID track with silicon detector hits. Muons from cosmic radiation are rejected by a requirement of a maximum distance of the muon track from the primary vertex of the hard interaction along the z axis.

Furthermore, muon candidates must be isolated from other high-p_{T} objects. Due to the considerable number of pile-up events present in the 2012 dataset, a so-called mini-isolation requirement is chosen. It normalises the allowed tracking activity in a cone around the muon candidate track to the p_{T} of the candidate,

$$I^{\mu}_{\mathrm{mini}} \coloneqq \sum_{i \in \{\mathrm{tracks}\}} \frac{p_{\mathrm{T}i}}{{p_{\mathrm{T}}}^{\mu}} < 0.05 \ . \tag{2.9}$$

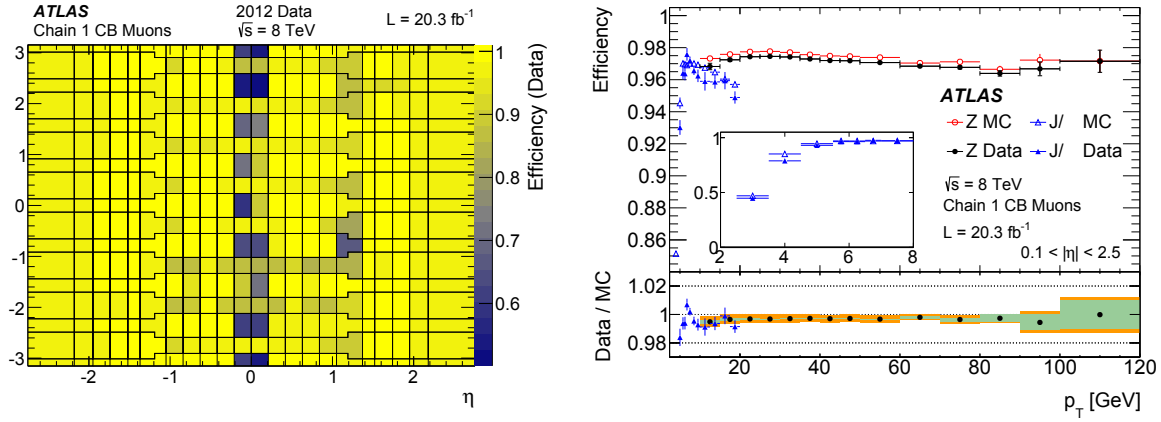
The sum involves tracks with $p_{\mathrm{T}i} > 1\,\mathrm{GeV}$ which meet additional quality criteria and are close to the muon candidate track, $\Delta R(i,\, \mu) < 10 \,/\, {p_{\mathrm{T}}}^{\mu}\,[\mathrm{GeV}]$. The muon track transverse momentum p^{μ}_{T} is excluded from this sum. Furthermore, the distance between muon candidates and jets must be sufficiently large, $\Delta R > 0.4$.

Figure 2.17 shows muon reconstruction efficiencies as function of p_{T}, η and φ, measured with tag-and-probe methods discussed at the end of this section. While the efficiencies are mostly close to one, reduced efficiencies are observed for central muons with $\eta \approx 0$ due to services of the ID and calorimeters installed in this region. The efficiency is also reduced in a solid angle at $1.1 < \eta < 1.3$ as some muon chambers were not yet installed in this region.

Momentum calibration

Muon momenta are calibrated based on samples of measured $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events. Additionally, $\Upsilon \to \mu\mu$ events are used for validation purposes. From these measurements, η and φ-dependent corrections of simulated muon transverse momenta regarding the momentum scale and resolution are derived. The momentum magnitude is corrected by a linear scaling while the resolution correction is carried out by means of momentum smearing in regions where the momentum spread is underestimated by the simulation. In this context several powers of muon p_{T} are multiplied with random numbers and different correction factors. The corrections are performed for the ID and the MS individually, while the combined muon p_{T} is given by an average value of the corrected ID and MS transverse momenta.

The correction factors are derived by fitting distributions of the di-muon masses $m^{\mathrm{ID}}_{\mu\mu}$ and $m^{\mathrm{MS}}_{\mu\mu}$, constructed from ID and MS tracks, respectively, to the data. As the MS p_{T} corrections are more complicated, the relative deviation of ID and MS p_{T} is used as an additional fit variable in this case. Accordingly, the ID corrections are derived first. The fit variables are separated into different p_{T} bins and template signal distributions are



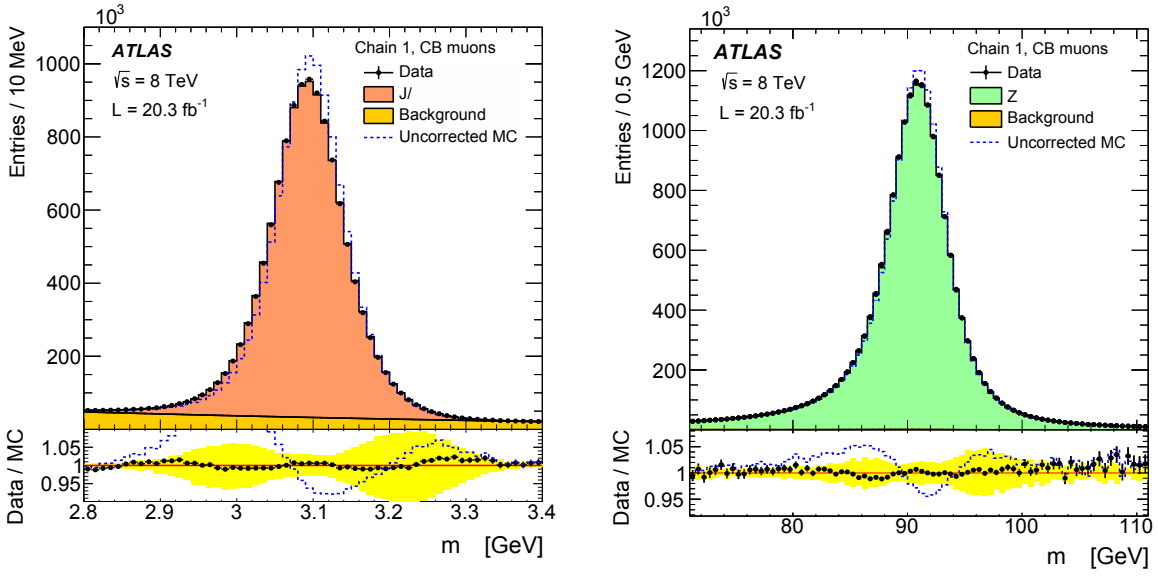
(a) (b)

Figure 2.17.: ATLAS muon reconstruction efficiency. In (a) measured reconstruction efficiencies are shown for combined muons as a function of their direction (η, φ) given $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. In (b) reconstruction efficiencies of combined muons are shown as a function of the muon p_{T}. While high-p_{T} muon efficiencies are measured with the help of $Z \to \mu\mu$ events, $J/\psi \to \mu\mu$ events are used in case of low p_{T} muons. The latter case is shown in the inlay of the figure. The lower panel compares efficiencies obtained from measurements and simulations. All efficiencies refer to muons reconstructed with the STACO algorithm instead of the Muid algorithm relevant to the measurement presented in this thesis. However, main characteristics of efficiencies are shared between the two types [154].

derived from simulations, while background contributions to the $J/\psi \to \mu\mu$ selection are estimated from data with the help of an analytical model. The momentum scale corrections resulting from the fits are on the order of 0.1% or smaller, while the resulting resolution corrections are on the order of 10%. Figure 2.18 shows the effect of the momentum corrections on di-muon invariant mass distributions regarding $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events.

Muon Trigger

The main single muon triggers used without prescales by ATLAS in 2012 are called `EF_mu24i_tight` and `EF_mu36_tight` [143]. Both triggers are seeded by the LVL-1 item MU15, where `15` indicates the p_{T} threshold. On LVL-2, an MS track needs to be reconstructed with $p_{\mathrm{T}} > 6\,\mathrm{GeV}$. This track must be matched to an ID track and the combined p_{T} must fulfill $p_{\mathrm{T}} > 22\,\mathrm{GeV}$. Finally, the Event Filter refines the momentum reconstruction. Aside from the different p_{T} thresholds of 24 GeV and 36 GeV, respectively, the two trigger chains differ by a loose isolation requirement which is applied by the



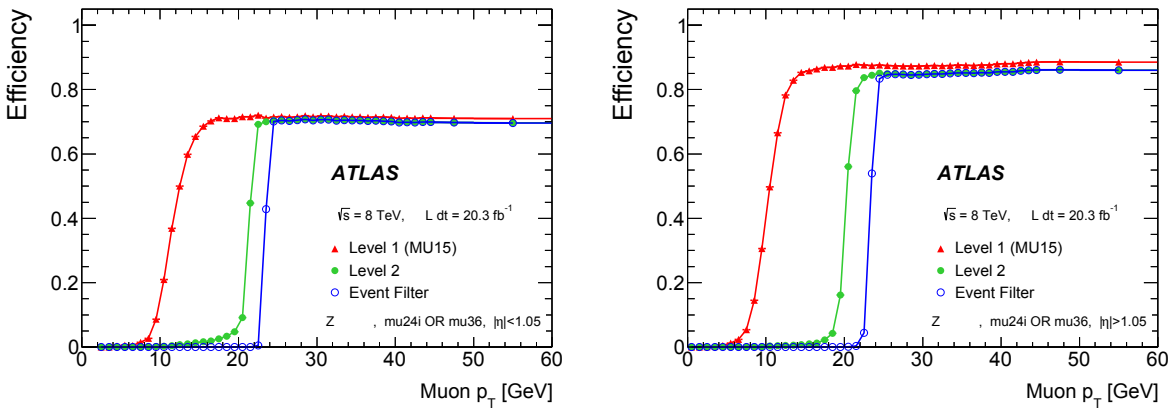
(a) (b)

Figure 2.18.: ATLAS muon momentum calibration. The results of the calibration are shown in terms of di-muon mass distrubutions for $J/\psi \to \mu\mu$ events (a) and $Z \to \mu\mu$ events (b). After corrections are applied, the simulated distributions match the data much better. The lower panels compare the measured distributions with the corresponding corrected and uncorrected simulations. The bands show the effect of systematic uncertainties on the momentum corrections. As in Fig. 2.17, results for combined muons reconstructed using the STACO algorithm are shown [154].

`EF_mu24i_tight` trigger. However, the mini-isolation requirement applied offline is tighter than this trigger isolation requirement. The resulting trigger efficiencies for the three trigger levels are shown in Fig. 2.19. The corresponding measurements are described below.

Efficiency measurements

The tag-and-probe method is used in order to measure muon reconstruction efficiencies and to match efficiencies in simulations to the data. In principle the procedure is similar to the electron case (Sec. 2.5.2), but for muons two independent momentum measurements are given by the ID and MS. Tracks reconstructed in one subdetector serve as probes while tags are searched for in the other subdetector. Combined muon reconstruction



(a) **(b)**

Figure 2.19.: ATLAS muon trigger efficiency. It is defined as the fraction of measured probe muons passing either the mu24i or mu36 trigger chain. The efficiencies are shown for each of the trigger levels separately and with respect to the probe muon p_{T} for the barrel region (a) and for the end-cap regions (b). The trigger efficiency is lower in the barrel region because of a reduced geometric coverage provided by the trigger chambers in this region as described in Sec. 2.2.4 [143].

efficiencies $\varepsilon(\text{reco})$ are given by:

$$\begin{aligned}\varepsilon(\text{reco}) &= \varepsilon(\text{reco}|\,\text{ID}) \cdot \varepsilon(\text{ID}) \\ &\approx \varepsilon(\text{reco}|\,\text{ID}) \cdot \varepsilon(\text{ID}|\,\text{MS}),\end{aligned}$$

where $\varepsilon(\text{reco}|\,\text{ID})$ is the efficiency to reconstruct a combined muon, given an ID track, and $\varepsilon(\text{ID})$ is the ID tracking efficiency. As the latter efficiency cannot be measured directly, it is replaced by the ID tracking efficiency, given an MS track, $\varepsilon(\text{ID}|\,\text{MS})$. This approximation is validated using a comparison of this tag-and-probe approach with an efficiency determination which is only based on simulations where generated muons serve as unbiased probes. The efficiencies agree with each other within less than 0.1% except for the region $|\eta| < 0.1$. Furthermore, this difference cancels to some extent regarding the SFs $\varepsilon^{\text{data}}(\text{reco})/\varepsilon^{\text{MC}}(\text{reco})$ which are used to correct acceptances of the simulations.

The SF measurements are mainly based on $Z \to \mu\mu$ events. The tag muon must be a combined muon which triggered the event readout. The probe muon must have an MS track concerning the measurement of $\varepsilon(\text{ID}|\,\text{MS})$. For the measurement of $\varepsilon(\text{reco}|\,\text{ID})$, on the other hand, the probe muon must have an ID track associated with an energy deposit in the calorimeter which corresponds to a minimum ionising particle. The efficiencies are given by the fraction of probes matched to reconstructed muons. The number of

background events in the selected sample is estimated from simulations or in a data-driven way and their contributions are subtracted from the overall measured numbers of probes and matches. Measurements of low-p_{T} efficiencies are based on $J/\psi \to \mu\mu$ events, where the numbers of probes and matches are estimated by fits of the di-muon mass.

Concerning trigger efficiencies, the tag-and-probe method is applied in a way similar to the reconstruction efficiency measurements. In addition, events containing muons where the readout was caused by missing transverse momentum triggers based only on calorimeter information are used. These are mainly events with W bosons and associated jets as well as top-quark events. They serve trigger efficiency measurements for high muon transverse momenta $p_{\mathrm{T}} > 100\,\mathrm{GeV}$. In addition, these selections allow for cross-checks of the data-MC SFs measured with the help of $Z \to \mu\mu$ events [143].

2.5.4. Jets

While in Sec. 1.2.2 the jet clustering of particles was discussed with regard to theoretical jet cross sections, the same clustering procedure can be applied to energy deposits in calorimeters in order to obtain experimental jet cross sections. For this purpose, the anti-k_{T} algorithm with a distance parameter of $R = 0.4$ is used in the measurement presented in this thesis. In this section the overall jet reconstruction procedure is discussed [155]. Once jets are identified by the clustering algorithm, their energy needs to be calibrated. In addition, quality criteria must be met. Also the identification of jets containing b-hadrons, called b-tagging, is discussed.

Reconstruction

The reconstruction of jets starts with the identification of clusters of energy deposits in the calorimeters, called topological clusters or topo-clusters [156]. These are groups of neighbouring calorimeter cells. Their grouping is seeded by energy deposits which are significantly greater than the expected noise contributed by electronics and pile-up events. Initially, energies are considered using the EM scale, which is appropriate for the reconstruction of EM showers. However, hadronic showers are more complicated. Aside from EM interactions, caused e.g. by $\pi_0 \to \gamma\gamma$ decays, they also involve nuclear interactions. In particular, a part of their energy is not deposited in the calorimeter sensors but is consumed e.g. by the binding energy of nucleons or neutrinos produced in

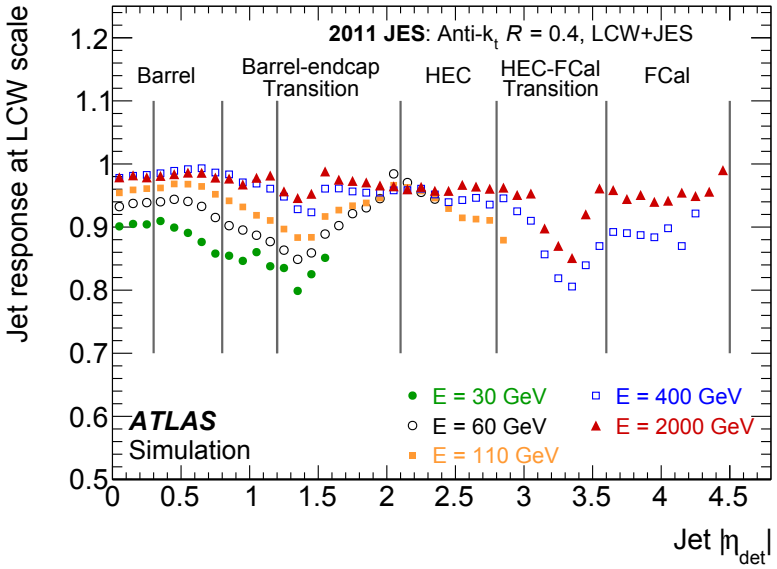


Figure 2.20.: ATLAS jet energy response before the final calibration. Calorimeter jets are reconstructed in simulated events using topo-clusters and the LCW scale. The energy response $\mathcal{R} = E_{\text{jet}}^{\text{reco}}/E_{\text{jet}}^{\text{truth}}$ is shown for different energies $E_{\text{jet}}^{\text{truth}}$ of jets formed from stable generated particles with which the calorimeter jets match and as a function of the pseudorapidity. For jets measured in experimental data, residual calibrations need to be applied additionally [155].

hadron decays. The impact of hadronic shower fluctuations is reduced by a dedicated weighting technique called local cell signal weighting (LCW). Here topo-clusters are classified as either EM or hadronic, based on properties like the energy density or the shower depth. Then the corresponding energy correction derived from simulations is applied to these clusters.

Once the LCW topo-clusters are identified, they are used as input to the anti-k_{T} jet clustering algorithm. The initial topo-clusters are treated as massless entities while in each merging step the four momenta of the merged clusters are added, which in turn introduces finite masses. Since for massive objects only rapidity differences rather than pseudo-rapidity differences are invariant with respect to longitudinal Lorentz boosts, rapidity differences are used in the distance measure of the jet algorithm, $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\varphi)^2}$ (Eq. 2.2).

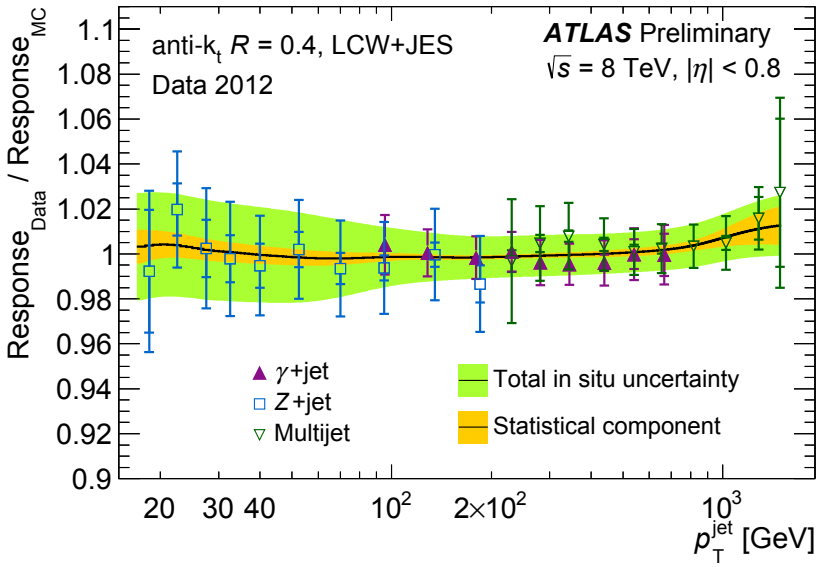


Figure 2.21.: ATLAS absolute jet energy in-situ calibration. The average jet response $\langle {p_T}^{jet}/{p_T}^{ref} \rangle$ measured in data is compared with the corresponding ratio obtained from simulations. Three methods are used, each of which employs different reference objects. Finally, the individual results are combined, which in turn results in the black line [139].

Energy calibration

Once the jet-algorithm has identified all jets based on the corrected topo-clusters, three steps are taken in order to match the jet energy scale to the one of jets reconstructed in simulations at the level of stable particles. First of all, jet energies are corrected for the offset caused by additional *pp* interactions in the bunch crossing of interest and by neighbouring bunch crossings. The corrections depend on the solid angle covered by the jet, the pseudorapidity of the jet, the measured number of primary vertices and the mean number of *pp* interactions during one bunch crossing. They are derived from simulations.

In a second step, jet energies and pseudorapidities are calibrated based on corrections derived from simulated jet events where reconstructed jets are compared with so-called truth jets based on simulated stable particles. At this point, the jet calibration is called "LCW+JES calibration". The energy response concerning such jets, $\mathcal{R} = E_{\text{jet}}^{\text{reco}}/E_{\text{jet}}^{\text{truth}}$, which is the ratio of the reconstructed calorimeter jet energy in simulations to the energy of matched jets of generated stable particles, is shown in Fig. 2.20.

Jets reconstructed in experimental data are further calibrated. Differences between measured and simulated jets are identified in order to take into account effects which are not sufficiently well modelled in the context of the LCW+JES calibration based only

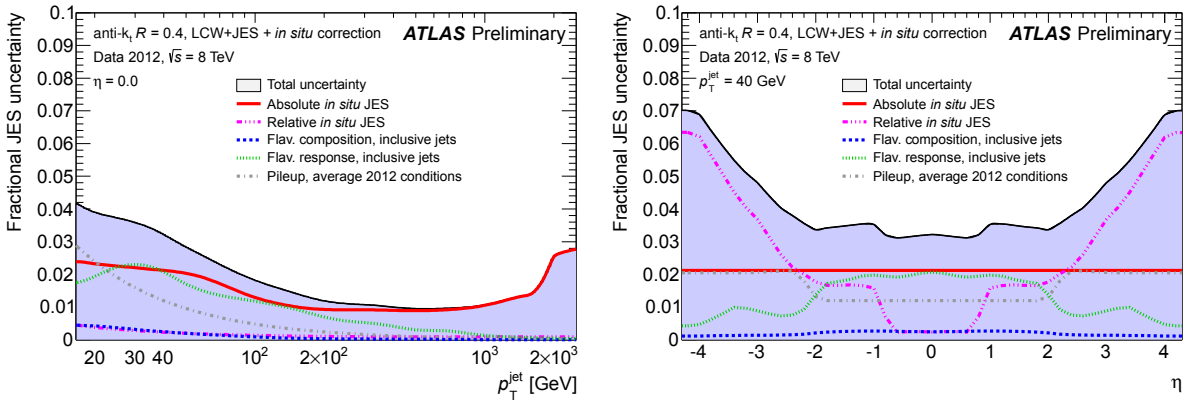
on simulated events. The jet momenta measured in data are corrected accordingly. As this procedure relates the measured jets to jets reconstructed in simulations, whereas the latter are related to jets of simulated stable particles by means of the LCW+JES calibration, these hadron level jets are the actual calibration reference.

While the absolute jet energy scale is explicitly determined for central jets, jet energies are calibrated relative to each other across different detector regions. This step is necessary due to the variance of employed calorimeter technologies and amounts of material. In the context of this so-called η-intercalibration, dijet events are used and transverse momentum correction factors are defined for different pseudorapidity regions. The p_{T} ratios of the two jets reconstructed in each event are collected in a histogram which is binned with respect to the pseudorapidities of the jets. Correction factors are derived from a fit to this histogram. Finally, these correction factors are rescaled such that the average corrections in the central region equal unity, as the η-intercalibration only aims for a relative calibration with respect to the peudorapidity η. This procedure is performed for dijet events in data and simulations. The ratios of the correction factors obtained from data and simulations provide corrections which are applied to measured jets.

The in-situ techniques applied for the residual calibration of the absolute jet energy scale in experimental data take advantage of momentum conservation in the transverse plane, where the momenta of a central jet and a well calibrated reference object must be approximately equal. Different techniques are used in order to determine the absolute energy scale correction factors to be applied to jets measured in experimental data,

$$\mathcal{R}({p_{\mathrm{T}}}^{\mathrm{jet}},\, \eta^{\mathrm{jet}}) = \frac{\langle {p_{\mathrm{T}}}^{\mathrm{jet}} / {p_{\mathrm{T}}}^{\mathrm{ref}} \rangle_{\mathrm{MC}}}{\langle {p_{\mathrm{T}}}^{\mathrm{jet}} / {p_{\mathrm{T}}}^{\mathrm{ref}} \rangle_{\mathrm{data}}} \,. \tag{2.10}$$

Z+jet events involve only small background contributions and are suited for the calibration of jets with relatively low p_{T}. Here events which involve an electron and a positron with an invariant mass close to the Z boson pole mass, and one high-p_{T} jet are selected. Due to effects like additional radiation, pile-up and the underlying event, the jet p_{T} is actually not required to be equal to the Z boson p_{T}. Instead, the ratio ${p_{\mathrm{T}}}^{\mathrm{jet}} / \left({p_{\mathrm{T}}}^{Z} \cdot |\cos \Delta\varphi\,(\mathrm{jet},\, Z)\,| \right)$ is used for comparisons of measurements with simulated events. The additional cosine of the azimuthal angle difference is used to reduce the impact of additional radiation as it reduces the reference p_{T} in case of an azimuthal imbalance.



(a) (b)

Figure 2.22.: ATLAS jet energy scale uncertainty. The relative jet energy scale uncertainty is shown for central jets as a function of the jet p_{T} (a) and for jets with $p_{\mathrm{T}} = 40\,\mathrm{GeV}$ as a function of pseudorapidity (b). Uncertainty components which depend on the event topology refer to inclusive dijet samples. The overall uncertainty is dominated by the residual in situ calibration [139].

Concerning harder jets with $p_{\mathrm{T_{jet}}} > 150\,\mathrm{GeV}$, γ+jet events are used because of their large statistics. The applied technique is based on the detector response to the total hadronic recoil. In the selected events with a photon and a jet, this recoil is supposed to balance the photon p_{T}. A detector response to the recoil other than unity results in a finite measured missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$. This variable, which is dicussed in the next section, is used for the calibration of the detector response.

For very high transverse momenta $p_{\mathrm{T}} > 800\,\mathrm{GeV}$, multijet events are used. There the hardest jet p_{T} is balanced against the p_{T} of a system of softer jets for which calibrations based on Z+jet or γ+jet calibrations are available. This method is iterated in order to use lower p_{T} jets with transverse momenta beyond the reach of Z+jet and γ+jet calibrations.

In Fig. 2.21, jet responses in data and in simulations are compared with each other for the different calibration approaches. Finally, the data-MC response ratios of these different methods are combined in terms of a weighted average taking into account the individual uncertainties. Furthermore, pile-up events must be taken into account in addition to the hard scattering process causing the jet. For this purpose, a pile-up p_{T} area density is introduced. Multiplied with the jet area, it provides the mean transverse momentum contributed to the jet by soft pp interactions. This contribution is subtracted

from the jet p_{T} [157]. Figure 2.22 shows examples of jet energy scale uncertainties resulting from the overall calibration procedure.

Energy resolution

The jet energy resolution is measured by means of the bisector method [158]. Dijet events are selected and the transverse momentum of the system of the two jets is considered. Finite energy resolutions are reflected by finite values of this momentum sum. Being a sum of two nearly opposite momenta, its components along the bisector of the transverse jet momenta on the one hand, and along the direction orthogonal to the former on the other hand, roughly parallel to the two-jet axis, depend differently on the resolution. The overall jet energy resolution is extracted with the help of the variances of these components. The results obtained from measured data are mostly in accordance with their simulated counterparts except for small differences present in particular regions of phase space. These differences as well as uncertainties affecting the precision of this comparison are covered by additional energy smearings which are considered as systematic uncertainties in measurements [159].

Additional quality criteria

In order to reject backgrounds like jet signatures originating from interactions of protons with residual gas present in the beam pipe, cosmic particle showers or hardware failures, additional quality criteria must be met, for instance a matching of particle tracks to the jet candidates. In order to reduce the impact of additional soft pp interactions occuring next to the hard scattering, a cut on the jet vertex fraction is applied. Taking into account all tracks associated with a jet, this variable is defined as the fraction of transverse momentum associated with the corresponding primary vertex,

$$\mathrm{JVF} = \frac{\sum\limits_{i \,\in\, \{\text{jet tracks}\} \,\cap\, \{\text{PV tracks}\}} p_{\mathrm{T}i}}{\sum\limits_{i \,\in\, \{\text{jet tracks}\}} p_{\mathrm{T}i}} \,. \tag{2.11}$$

A typical requirement is $\mathrm{JVF} > 0.5$. Uncertainties due to discrepancies of the JVF cut efficiency between data and simulations are estimtaded by means of corresponding cut variations. These variations are derived from dedicated studies of events with $Z \to ee$

decays and a jet. Here the Z boson momentum balances the momentum of the jet, so that a pure sample of jets originating from hard interactions is obtained.

Once all jets are reconstructed, overlaps between jets and electrons need to be removed. For this purpose, one jet candidate reconstructed closest to a reconstructed electron is discarded if its distance to the electron is small, namely $\Delta R < 0.2$.

Identification of b-jets

Jets which involve hadrons containing b-quarks are called b-jets. Their identification, known as b-tagging, is an important technique for many measurements at ATLAS, including the single top-quark measurement presented in this thesis due to the expected top-quark decays into b-quarks. The relatively long lifetime of the resulting b-hadrons and their large boosts resulting from the decay of the massive top-quark lead to vertices which are displaced by typically a couple of millimeters with respect to the primary vertex of the hard pp interaction.

The tagging algorithm relevant to this thesis combines several observables related to these secondary vertices (SVs). For instance, these are transverse and longitudinal impact parameters of tracks which are associated with the jet, an illustration of which is shown in Fig. 2.23. Other employed observables are based on explicit identifications of SVs by means of vertex fits. These observables are, among others, the distance between the SV and PV, the invariant mass of all tracks associated with the SV and the energy fraction of the sum of the SV tracks with respect to the set of all tracks associated with the given jet. The chosen observables are used as input to different tagging algorithms which provide a b-tag weight for the jet. Finally, these weights are used as input for a combined tagging algorithm called MV1, which is a neural network implemented with the help of the TMVA package [160]. The output of the MV1 algorithm is a single weight which allows one to identify b-jets.

In measurements, b-tagging is performed by means of a cut on the b-tag weight. Typically a working point corresponding to a b-jet selection efficiency of 70% is chosen. In order to ensure an appropriate description of measured data by simulations, the b-tagger needs to be calibrated. For this purpose the efficiency of the b-tagging cut needs to be determined in measured data and in simulations, so that data-MC scale factors can be computed. Clearly, for this calibration a pure sample of b-jets is desired. One suitable source of b-jets at ATLAS are top-quark pair production events as they are abundant and

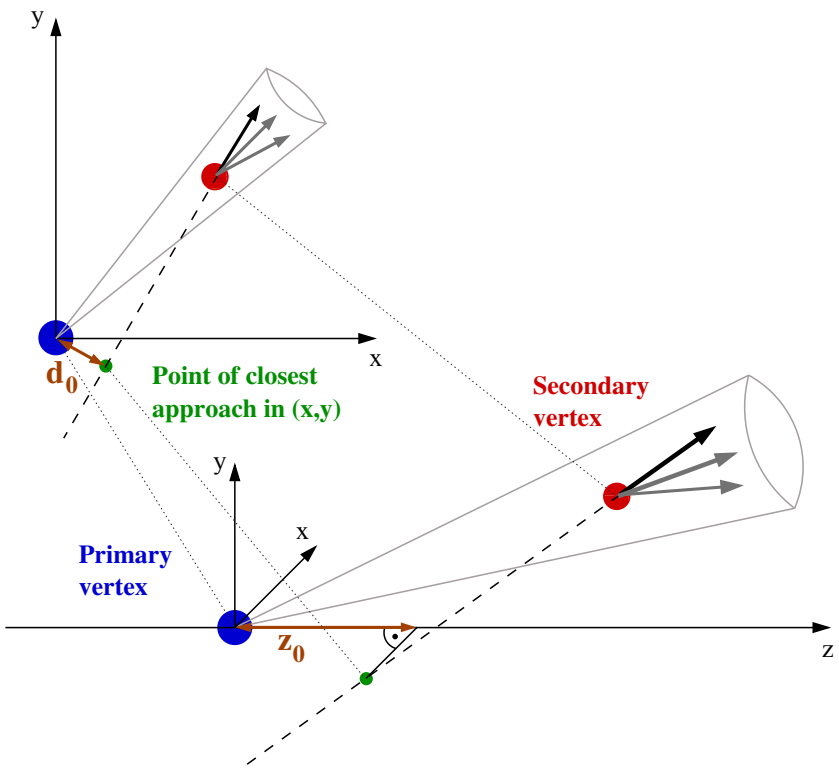
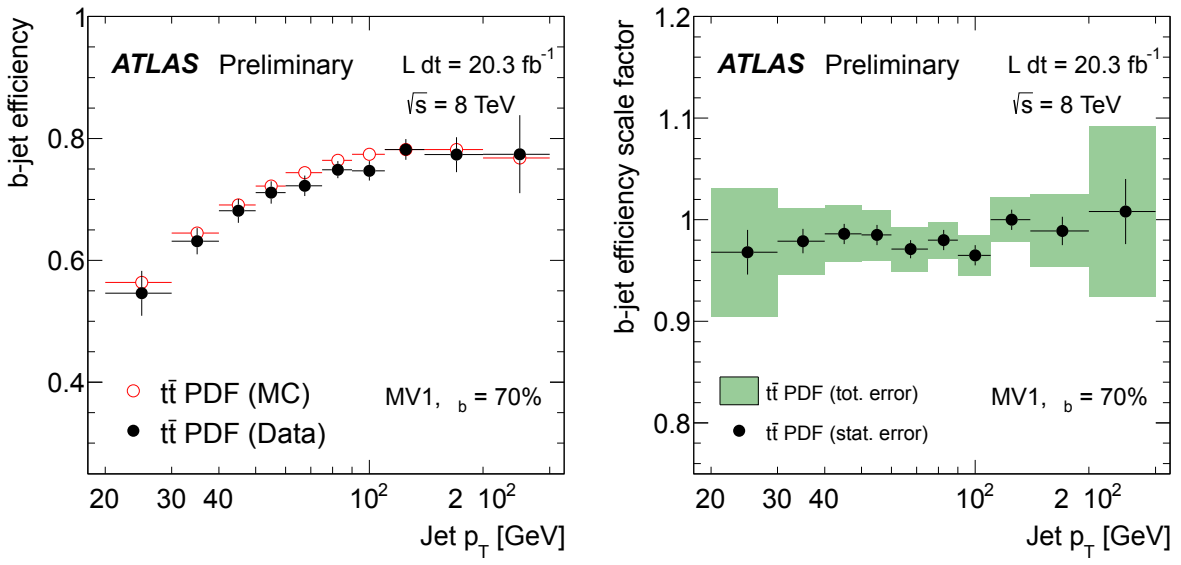


Figure 2.23.: Impact parameters of charged particle tracks. Within b-jets, decays of b-hadrons cause secondary vertices which are displaced with respect to the primary vertex of the hard interaction. Once the tracks originating from these secondary vertices are extrapolated to the beam spot, deviations from the primary vertex show up. Using the projection of a track onto the transverse plane, the point of closest approach provides the transverse impact parameter d_0, as shown in the upper illustration. The longitudinal deviation between the corresponding point of the three-dimensional track and the primary vertex provides the longitudinal impact parameter z_0, as shown in the lower illustration. These parameters are used for the tagging of b-jets.



(a) (b)

Figure 2.24.: ATLAS b-tagging calibration. In (a) the b-tagging efficiency of the MV1 tagging algorithm at a 70% efficiency working point in data and simulations is shown. The results are based on a set of di-leptonic $t\bar{t}$ events. From the ratio of the two efficiencies, data/MC scale factors are derived (b). While the scale factors mostly cover $\varepsilon_{\text{data}}/\varepsilon_{\text{MC}} = 1$, their uncertainties are an important input for physics anlyses [161].

contain b-jets due to the top-quark decays into b-quarks. However, in such calibrations it is assumed that the CKM matrix element $|V_{tb}|$ equals one, as basically predicted by the SM.

In the following a calibration based on di-leptonic $t\bar{t}$ events is briefly described, as this one is relevant to the single top-quark measurement presented in the next chapter. Here, events with two oppositely charged electrons or muons and two jets are selected. In the context of a so-called "combinatorial likelihood method", the b-tagging efficiency of a given weight cut is the parameter of interest in an unbinned maximum likelihood fit. The likelihood function for one event depends on the transverse momenta p_{T} and b-tag weights w of the two jets,

$$L(p_{\text{T}1},\, p_{\text{T}2},\, w_1,\, w_2) = \frac{1}{2} \sum_{i,\,k\,\in\,\{1,\,2\}}^{i\neq k} \sum_{\alpha,\,\beta\,\in\,\{b,\,j\}} f_{\alpha\beta}\mathscr{P}_{\alpha\beta}(p_{\text{T}i},\, p_{\text{T}k})\mathscr{P}_{\alpha}(w_i|\, p_{\text{T}i})\mathscr{P}_{\beta}(w_k|\, p_{\text{T}k})\ . \tag{2.12}$$

Here $f_{\alpha\beta}$ denotes the jet flavour fractions for light jets j and b-jets b, while $\mathscr{P}_{\alpha\beta}(p_{\mathrm{T}i}, p_{\mathrm{T}k})$ denotes the probability density of the two transverse momenta of the jets given a flavour combination $\alpha\beta$. These entities are derived from simulations. $\mathscr{P}_{\alpha}(w_i|\, p_{\mathrm{T}i})$ denotes the probability density of the tag weight for a jet flavour α, given the jet p_{T}. While for light jets, $\alpha = j$, this density is derived from simulations as well, the b-jet tag weight density, $\alpha = b$, contains the fit parameters. It is approximated by a step function and implemented using histograms with two weight bins. The upper bin beyond the weight cut corresponds to the desired tagging efficiency,

$$\varepsilon_b(p_{\mathrm{T}}) = \int_{w_{\mathrm{cut}}}^{1} \mathrm{d}w\, \mathscr{P}_b(w|\, p_{\mathrm{T}}) \ . \tag{2.13}$$

These efficiencies result from the maximum likelihood fit using all selected events. Ranges of transverse momentum are distinguished in order to obtain p_{T}-dependent b-tag efficiencies. The procedure is cross-checked on the basis of simulations. No significant bias is found. Examples of derived efficiencies and data-MC scale factors are shown in Fig. 2.24, provided by Ref. [161], where more details are discussed, including the estimation of background contributions.

Apart from the tagging of actual b-jets, the response of the tagging algorithm to charm and light-flavour jets is calibrated as well. Concerning charm hadron jets, this is accomplished on the basis of events which contain jets including D^* mesons. The latter can be identified efficiently due to distinct decay topologies. The response to light flavour jets, on the other hand, is calibrated using multijet events [162].

2.5.5. Missing transverse momentum

While all physics objects discussed so far can be reconstructed efficiently, this is not the case for neutrinos or possible other, yet undiscovered weakly interacting particles. The corresponding interaction cross sections are way too small to allow for a reliable detection of a single weakly interacting particle. However, the presence of such particles in an event can still be figured out with the help of momentum conservation in the transverse plane. Since the transverse momenta of the colliding partons are known to be small, imbalances of the transverse momentum in the final state, called missing transverse momentum or energy, denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$, are associated with weakly interacting particles produced in the collisions. The $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation involves energy deposits in the calorimeter and muon

momenta [163]. Calibrated physics objects are used. Calorimeter cells and tracks which are not associated with reconstructed jets or leptons enter as well. Using only objects which are relevant to the single top-quark measurement presented in this thesis, $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as follows,

$$\begin{aligned} E_{x,y}^{\mathrm{miss}} &= E_{x,y}^{\mathrm{miss},e} + E_{x,y}^{\mathrm{miss,jets}} + E_{x,y}^{\mathrm{miss,soft}} + E_{x,y}^{\mathrm{miss},\mu} \ , \\ E_{\mathrm{T}}^{\mathrm{miss}} &= \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \ . \end{aligned} \tag{2.14}$$

The contributions of electrons, jets, additional soft energy deposits and muons are negative energy or momentum sums of the corresponding objects, projected onto the x and y directions,

$$\begin{aligned} E_x^{\mathrm{miss,X}} &= -\sum_{i \in \{X\}} E_i \sin\vartheta_i \cos\varphi_i \ , \\ E_y^{\mathrm{miss,X}} &= -\sum_{i \in \{X\}} E_i \sin\vartheta_i \sin\varphi_i \ . \end{aligned} \tag{2.15}$$

The sums refer to sets of objects of one kind. If combined muon momenta are used, the corresponding energy deposit in the calorimeter is subtracted in order to avoid a double counting. Jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ enter the jet contribution $E_{x,y}^{\mathrm{miss,jets}}$. The soft term $E_{x,y}^{\mathrm{miss,soft}}$ is made of topological clusters where only significant energy deposits are taken into account in order to suppress contributions caused by noise. Clusters are calibrated using the LCW scheme. Tracks which are not associated with high-p_{T} objects also enter the soft energy term in order to recover contributions from low-p_{T} particles. Overlaps between tracks and topo-clusters are removed. Uncertainties related to the scale and the resolution of the soft term are estimated on the basis of $Z \to \mu\mu$ events which do not contain a genuine $E_{\mathrm{T}}^{\mathrm{miss}}$.

Chapter 3.

Single top-quark measurement

"Please please please
No more remedies
My method is uncertain
It's a mess, but it's working
And maybe, if you want to try it out
You won't like it so, when you're crying out:
Give us something familiar; something similar
To what we know already
That will keep us steady
Steady
Steady
Steady
Steady going nowhere"
— Fiona Apple, *Please Please Please*

3.1. Analysis basics

The measurement presented in the following makes use of the ATLAS detector introduced in chapter 2 to test the theoretical prediction of the s-channel single top-quark production cross section discussed in chapter 1. The employed data set was recorded in 2012. A suitable selection of collision events is analysed with the method of total event likelihoods, which separates the s-channel signal from various background processes. This separation allows for a powerful search for the signal on the basis of an appropriate statistical model.

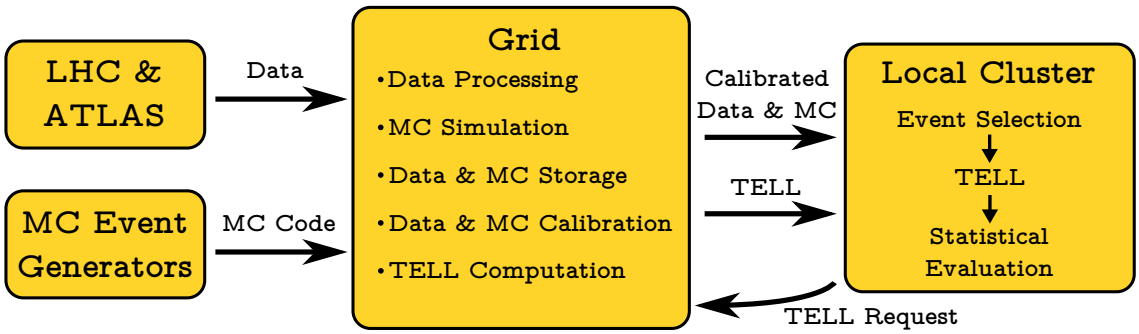


Figure 3.1.: Analysis data flow. LHC collision data recorded with the ATLAS detector is processed at the CERN computing centre, which is part of the Grid. MC simulation samples are processed on the Grid as well. Subsequently, the physics analysis is performed an a local computer cluster.

In the present section the employed data set, simulation samples, physics objects and collision events are discussed. The infrastructure used to carry out the analysis and the flow of data are illustrated in Fig. 3.1. LHC collision data is recorded by ATLAS and subsequently sent to the Worldwide LHC Computing Grid for processing purposes. In order to be able to interpret the data, MC event generators provide event simulation code which is also processed on the Grid. In addition, it provides mass storage capacity for data and MC samples. The raw data and MC events need to be calibrated, which is done on the Grid as well. Calibrated events are transferred to a local computer cluster at Humboldt-Universität zu Berlin. There events are selected with regard to the desired measurement. Subsequently, total event likelihoods, or TELLs, are computed. This task is shared between the local cluster and the Grid. Finally, a statistical evaluation is performed on the local cluster, which leads to the measurement results.

3.1.1. Data set

The complete set of pp collisions at a centre-of-mass energy of 8 TeV provided by the LHC and recorded by the ATLAS detector in 2012 is considered. The data is filtered given the requirement of proper data taking conditions of all subdetectors which are discussed above in Sec. 2.2. This is a necessary demand for the desired measurement to be efficient due to the diversity of physics objects resulting from top-quark decays, the reconstruction of which requires an operation of all major components of the ATLAS detector. Given this criterion, the data set corresponds to an integrated luminosity of

$20.3\,\mathrm{fb}^{-1}$, where an uncertainty of 2.8% needs to be taken into account. The luminosity calibration is based on dedicated beam separation scans as outlined in Sec. 2.3.

Concerning the readout and selection of collision events, the single electron and single muon triggers discussed in the sections 2.5.2 and 2.5.3 are used. These are the triggers with the lowest p_{T} thresholds among all unprescaled triggers of their kind used by ATLAS in 2012. Overlaps of the two trigger decisions are allowed. The choice of these two items is driven by the aim to focus on top-quark decays involving electrons or muons in order to avoid large background contaminations given by multijet events of strong interactions. Furthermore, the fact that these triggers do not involve any prescale allows one to gather large samples of top-quark events.

3.1.2. Modelling of processes

In order to be able to identify the s-channel single top-quark production signal despite the occurrence of large background contributions, the measurement is performed on a suitable subset of the available data, similarly to the analyses discussed in Sec. 1.3.4. Selected events contain signatures of leptonic W boson decays as well as b-tagged jets. Details of the event selection follow in Sec. 3.1.3. There are several processes beyond the s-channel signal which contribute to this event sample in large amounts. Furthermore, the event topologies are rather complex. Hence a thorough understanding of their kinematics is mandatory for the measurement to be sensitive to the signal. For these reasons, extensive use is made of MC event generators. They allow one to interface the experimental data with theoretical predictions. In concrete terms, the topology of signal events is assumed to follow the SM prediction as implemented in an appropriate event generator. Only the normalisation of the signal process is subject to the measurement. The modelling of background processes is entirely constrained by SM predictions. However, there is one exception to this choice, namely background events caused by fakes of prompt leptons. This process can hardly be described by event generators in an efficient manner. Therefore its contribution is estimated from experimental data directly. In the following, all of the relevant backgrounds are introduced first, followed by a discussion of the modelling of scattering processes by means of MC event generators. The estimation of the lepton fake contribution is presented subsequently.

Background processes

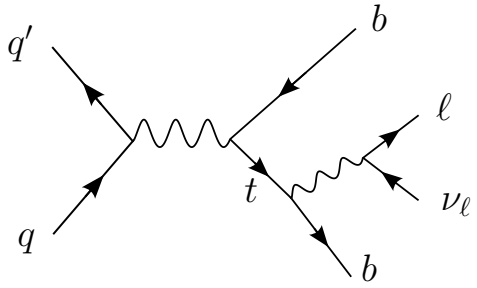
Scattering processes are considered to be relevant backgrounds if they can result in the event topology defined by the s-channel single top-quark production signal. This criterion does not refer to the set of all final state particles but rather to the set of reconstructed objects. Due to limited object reconstruction efficiencies and finite mis-identification rates, some processes can contribute to the selected sample of events even though they do not quite lead to the signal signature on the particle level. In the following introduction of the backgrounds, only generic event properties are discussed, while higher order effects can be neglected. A collection of selected leading order Feynman diagrams of the relevant processes is given in Fig. 3.2.

single top-quark production, t-channel: Most of the t-channel events are quite different from the s-channel signal. The top-quark is usually accompanied by a jet induced by a light quark. This jet can be missed in the event selection. Aside from the b-quark resulting from the top-quark decay, there is a second one involved in a splitting of an initial state gluon. However, jets induced by this b-quark seldom reach high p_{T}. Still, events with two b-tagged jets occur. Due to the relatively high cross section of the t-channel, these events must be taken into account.

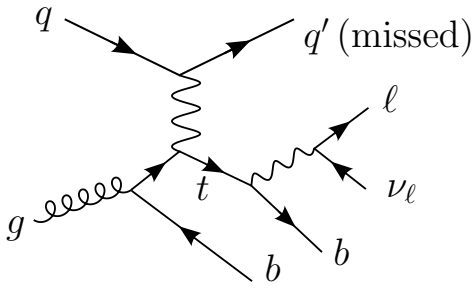
Single top-quark Wt production: Events of Wt associated production contain up to two charged leptons, depending on the decays of the two W bosons involved. Apart from the di-lepton case, events contain light jets resulting from W boson decays into quarks. Again, a final state b-quark can result from an initial gluon splitting, aside from the top-quark decay. Given the imperfection of the event reconstruction, Wt production must be taken into account as a background.

Top-quark pair production: Final states with pairs of top-quarks are similar to those of Wt single top-quark production except for the frequent reconstruction of two b-tagged jets caused by the two top-quark decays. Since this mode comes with the largest cross section of all top-quark production processes, it is the dominant background of the s-channel search.

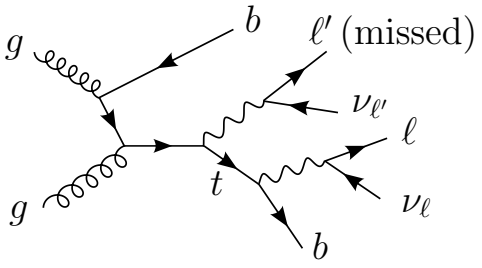
W+jets production: The production of a W boson can be accompanied by quarks produced via strong interactions. In particular, this includes the production of b-quarks. Hence the final state of the signal process can result. Jets induced by lighter quarks contribute as well since the mis-identification rate of the b-tagging is finite.



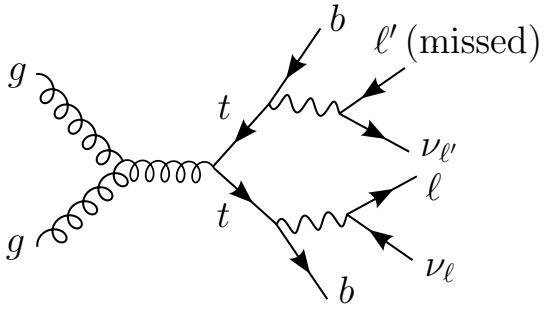
(a) single top-quark s-channel



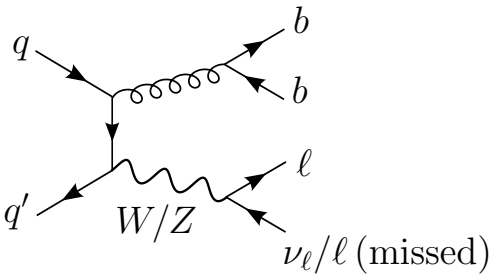
(b) single top-quark t-channel



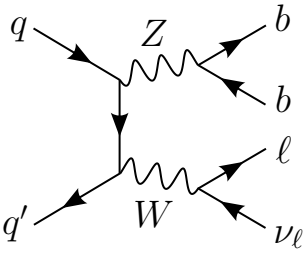
(c) Wt production



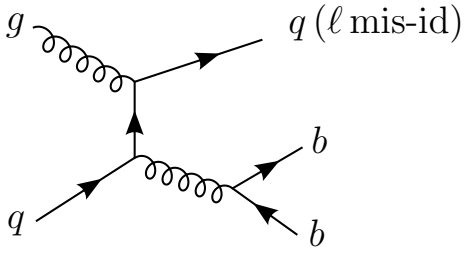
(d) top-quark pair production



(e) W/Z+jets



(f) Di-boson production



(g) Multijets

Figure 3.2.: Examples of Feynman diagrams of signal and background processes to be considered in the search for s-channel single top-quark production. The generic signature of the signal is given by two b-tagged jets, a charged lepton and large $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by a neutrino. The same signature can result from various background processes, in particular due to limited reconstruction efficiencies and finite mis-identification rates.

Z+jets production: The previous remarks concerning W+jets production also matter in case of Z+jets production. However, the relevant decays of the Z boson involve two charged leptons rather than a charged lepton and a neutrino. Furthermore, Z boson production cross sections are smaller. Hence the Z+jets processes are a background of minor importance.

Di-boson production: The production of two massive vector bosons WW, WZ or ZZ needs to be taken into account as well, due to various decay modes of the bosons and possible quarks being produced additionally via strong interactions.

Fakes of prompt leptons: The electron and muon identifications discussed in Sec. 2.5.2 and 2.5.3 result in high purities. Still, selected high-p_{T} electrons and muons do not necessarily originate from W boson decays. First of all, decays of hadrons into leptons can contribute to the selection if the isolation requirements fail. In addition, photon conversions $\gamma \to e^+e^-$ give rise to secondary electrons and positrons as well. Secondly, the lepton signatures can be faked by hadrons. Because of the largeness of jet cross sections at the LHC, these processes need to be taken into account.

Simulation samples

The event simulation involves the generators discussed in Sec. 1.4. They provide events consisting of particles which live long enough to reach the ATLAS detector. The detector response is simulated up to the output of readout electronics as described in Sec. 2.4. Most of the employed MC samples make use of the full detector simulation based on GEANT4, while the fast simulation program ATLFAST-II is used in order to process several samples needed for estimations of modelling uncertainties.

The normalisation of the MC samples must be adjusted with regard to the integrated luminosity of the data set in use and the total cross section of the scattering process in question. For this purpose, the luminosity of each bare sample is defined as follows,

$$\mathcal{L}_{\mathrm{MC}}^{\mathrm{bare}} = \frac{N_{\mathrm{MC}}^{\mathrm{bare}}}{\sigma_{\mathrm{MC}}^{\mathrm{bare}}} \ , \tag{3.1}$$

where $\sigma_{\mathrm{MC}}^{\mathrm{bare}}$ denotes the cross section resulting from the hard scattering simulation, which has LO or NLO accuracy, depending on the employed event generator. The sum of the weights of generated events is denoted by $N_{\mathrm{MC}}^{\mathrm{bare}}$. Each sample is reweighted in order to

fit its number of events to the event count expected in the measured data set,

$$\begin{aligned} N_{\mathrm{MC}}^{\mathrm{norm}} &\stackrel{!}{=} N_{\mathrm{data}} \\ &= \mathcal{L}_{\mathrm{data}} \sigma^{(\mathrm{N})\mathrm{NLO}} \\ &= \frac{\mathcal{L}_{\mathrm{data}}}{\mathcal{L}_{\mathrm{MC}}^{\mathrm{bare}}} \frac{\sigma^{(\mathrm{N})\mathrm{NLO}}}{\sigma_{\mathrm{MC}}^{\mathrm{bare}}} \mathcal{L}_{\mathrm{MC}}^{\mathrm{bare}} \sigma_{\mathrm{MC}}^{\mathrm{bare}} \\ &= \frac{\mathcal{L}_{\mathrm{data}}}{\mathcal{L}_{\mathrm{MC}}^{\mathrm{bare}}} \frac{\sigma^{(\mathrm{N})\mathrm{NLO}}}{\sigma_{\mathrm{MC}}^{\mathrm{bare}}} N_{\mathrm{MC}}^{\mathrm{bare}} \ . \end{aligned} \tag{3.2}$$

Hence, two weighting factors enter, namely the ratio of data over bare MC luminosities as well as the ratio of the most precise total cross section prediction available over the bare MC cross section. The cross section ratio is usually called the k-factor of the sample. The use of global k-factors ignores the fact that higher order corrections affect different regions of phase space differently. However, most of the MC event generators employed in the measurement provide NLO accuracy inherently. An overview of these generators is given above in Tab. 1.3. Additionally, QED radiation and τ-lepton decays are modelled by the Photos and Tauola MC event generators, respectively [164, 165]. Tables of detailed information on the corresponding MC sample production are given in App. A.

The modelling of all processes involving top-quarks is based on the Powheg-Box generator which simulates the hard scattering at NLO accuracy. Subsequent simulation steps are handled by Pythia 6. In order to study the extent to which the MC event generator choice affects the measurement results, alternative MC samples are produced using the aMC@NLO generator in case of s and t-channel single top-quark production, and MC@NLO in case of Wt and top-quark pair production. Both of these generators are interfaced with Herwig. In case of the s-channel signal, dedicated samples produced with the Powheg-Box and Pythia 6 generators using different renormalisation and factorization scales serve for an estimation of the corresponding uncertainties. The same is true in case of top-quark pair production, where in addition a damping parameter affecting the hardness of QCD radiation is varied. In all of these samples, a top-quark mass of 172.5 GeV is chosen. The uncertainty of the top-quark mass is particularly small, as discussed in Sec. 1.3.4. Thus it can be neglected in the present analysis given the level of measurement precision to be discussed below.

Special care is taken of associated Wt single top-quark production which interferes with the pair production of top-quarks as mentioned in Sec. 1.3.2. Concerning the event generation, two approaches have been developed to define electroweak Wt production up to NLO accuracy, namely the diagram removal (DR) and the diagram subtraction

(DS) schemes [96]. In the DR scheme, diagrams including two top-quark propagators are neglected. Hence it violates gauge invariance, but this shortcoming turns out to be insignificant in practice. In the DS scheme, a suitable term is subtracted at the cross section level. The difference of Wt cross sections calculated in the two schemes involves the interference term of Wt and $t\bar{t}$ production. In the present analysis, event samples generated in both schemes are used and compared with each other. Differences between the two schemes turn out to be insignificant. Furthermore, Wt production contributes only a small fraction to the selected set of events. In consequence, the treatment of the interference term does not affect the measurement.

The modelling of W+jets, Z+jets and di-boson events relies on the generators SHERPA 1.4.1, ALPGEN interfaced with PYTHIA 6 and HERWIG, respectively. These predictions provide leading order accuracy.

In accordance with the recommendations of the LHC Top Working Group, the s and t-channel single top-quark samples are normalised to the total cross sections at NLO accuracy discussed in Sec. 1.3.4, while the Wt production cross section is normalised to an approximate NNLO prediction [71, 72]. The top-quark pair production samples are normalised to the NNLO prediction discussed in Sec. 1.3.2. The single and di-boson samples are normalised to NNLO and NLO predictions, respectively [166, 167].

Lepton fake estimate

False identifications of prompt charged leptons have small probabilities to occur given any collision event. On the other hand, multijet events come with high rates. In these events fakes of prompt charged leptons can be caused by mis-identifications of hadrons or by barely isolated charged leptons originating from hadron decays. Therefore these fakes must be taken into account. Since simulations of lepton fakes are relatively inefficient, an estimation based on experimental data is used.

The basic idea of this approach is to derive the number of selected events which contain these fakes with the help of a superset of events based on a loosened lepton requirement [168]. Dedicated measurements of the probabilities of real and fake prompt leptons to pass the tight lepton selection given their presence in the loose selection sample allow for this derivation. Denoting these efficiencies by $\varepsilon_{\text{real}} = N_{\text{real}}^{\text{tight}}/N_{\text{real}}^{\text{loose}}$ and $\varepsilon_{\text{fake}} = N_{\text{fake}}^{\text{tight}}/N_{\text{fake}}^{\text{loose}}$, respectively, the number of events contained in the loose and tight

selections is given by

$$\begin{aligned} N^{\text{loose}} &= N^{\text{loose}}_{\text{real}} + N^{\text{loose}}_{\text{fake}} \ , \\ N^{\text{tight}} &= N^{\text{tight}}_{\text{real}} + N^{\text{tight}}_{\text{fake}} \\ &= \varepsilon_{\text{real}} N^{\text{loose}}_{\text{real}} + \varepsilon_{\text{fake}} N^{\text{loose}}_{\text{fake}} \ . \end{aligned} \tag{3.3}$$

It is assumed that only one lepton needs to be considered. This is true in case of the present measurement which deals with single lepton events. The system of equations can be solved for the desired number of fake leptons contained in the tight selection sample,

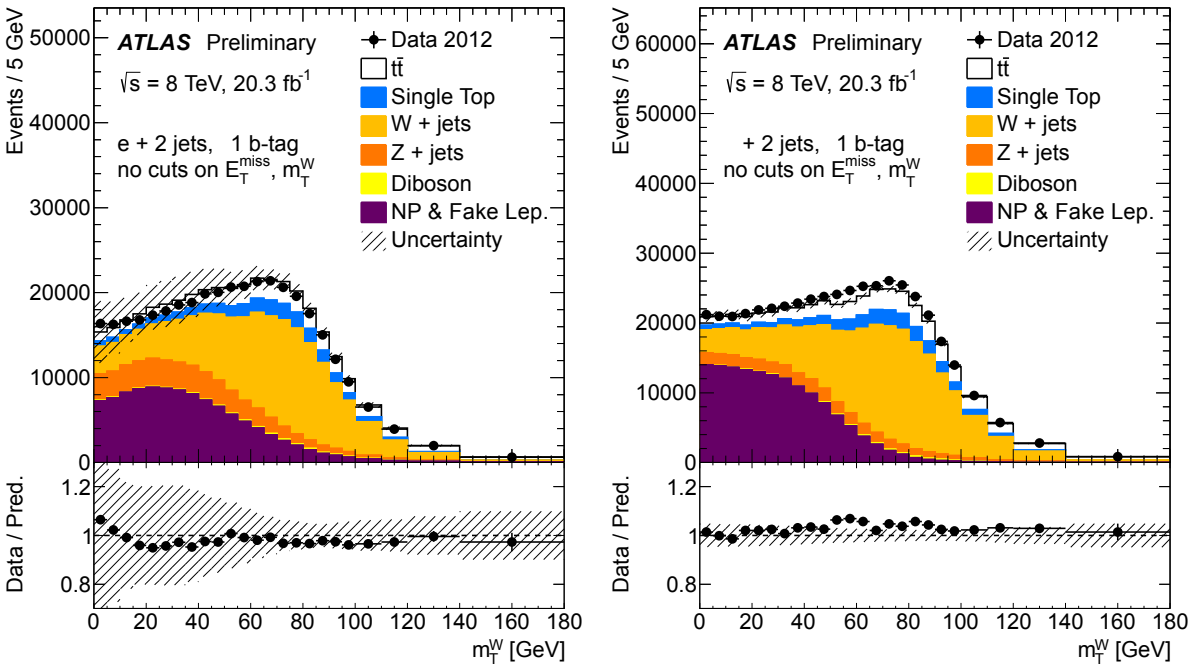
$$N^{\text{tight}}_{\text{fake}} = \frac{\varepsilon_{\text{fake}}}{\varepsilon_{\text{real}} - \varepsilon_{\text{fake}}} \left(\varepsilon_{\text{real}} N^{\text{loose}} - N^{\text{tight}} \right) \ . \tag{3.4}$$

The procedure is usually called the *matrix method*, which refers to this inversion of linear equations. In case of a simple counting experiment, equation 3.4 is sufficient. If kinematic distributions are taken into account, as in the case of the present measurement, the efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ must be determined as functions of the event kinematics. Then, Eq. 3.4 is applied for each bin of a histogram of a kinematic distribution of interest. In practice, this procedure is performed efficiently with the help of weights which are assigned to each event. For any bin k of any histogram

$$\begin{aligned} N^{\text{tight},\,k}_{\text{fake}} &= \frac{\varepsilon^k_{\text{fake}}}{\varepsilon^k_{\text{real}} - \varepsilon^k_{\text{fake}}} \left(\varepsilon^k_{\text{real}} N^{\text{loose},\,k} - N^{\text{tight},\,k} \right) \\ &= \sum_{i \,\in\, \{\text{loose}\} \setminus \{\text{tight}\} \cap \{\text{bin } k\}} \frac{\varepsilon^i_{\text{fake}}}{\varepsilon^i_{\text{real}} - \varepsilon^i_{\text{fake}}} \varepsilon^i_{\text{real}} \ + \sum_{i \,\in\, \{\text{tight}\} \cap \{\text{bin } k\}} \frac{\varepsilon^i_{\text{fake}}}{\varepsilon^i_{\text{real}} - \varepsilon^i_{\text{fake}}} \left(\varepsilon^i_{\text{real}} - 1 \right) \ , \end{aligned} \tag{3.5}$$

where the sums refer to the events of the loose and tight selections. The summands are assigned as weights to the selected events accordingly, so that Eq. 3.5 can be applied easily with regard to any kinematic distribution.

The efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are determined with respect to the tight electron and muon selections applied in the physics analysis itself, outlined above in Sec. 2.5.2 and 2.5.3. This includes the choice of triggers. The loose selection criteria involve a couple of relaxations. In particular isolation requirements are dropped. Concerning the employed event selection, the real lepton efficiencies $\varepsilon_{\text{real}}$ are measured in di-lepton events, where the tag-and-probe strategy introduced in Sec. 2.5.2 is used. Leptonic decays of Z bosons provide prompt electron and muon samples of high purity. The fake efficiencies $\varepsilon_{\text{fake}}$ are measured in events with high fractions of fakes of prompt leptons in both the loose and



(a) **(b)**

Figure 3.3.: ATLAS prompt lepton fake estimation using the matrix method. Selected events contain reconstructed electrons (a) or muons (b) and two jets, at least one of which is b-tagged. The distribution of the transverse W boson mass in measured data is compared with predictions given by simulations and the prompt lepton fake estimates. The shaded areas show the uncertainties of these estimates [168].

tight selection. These event samples are obtained by requiring low missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ and low transverse masses m_{T}^{W} given by

$$\begin{aligned} m_{\mathrm{T}}^{W} &= \sqrt{(p_{\mathrm{T}}^{\ell} + E_{\mathrm{T}}^{\mathrm{miss}})^2} \\ &= \sqrt{2|\mathbf{p}_{\mathrm{T}}^{\ell}||\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}| \left(1 - \cos\Delta\varphi(\mathbf{p}_{\mathrm{T}}^{\ell}, \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})\right)} \ . \end{aligned} \tag{3.6}$$

This is a measure of the likelihood of the charged lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ to originate from a W boson decay. Multijet events, on the other hand, usually lead to small values of m_{T}^{W}. The same is true for $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore events with $m_{\mathrm{T}}^{W} < 20\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} < 60\,\mathrm{GeV}$ are used to determine $\varepsilon_{\mathrm{fake}}$ in case of electrons. Muons, on the other hand, are unlikely to be faked by other particles. Fakes of prompt muons are mainly caused by secondary muons produced in decays of hadrons. Accordingly, their fake efficiencies are determined in events with a muon of large track impact parameter $d_0/\sigma_{d_0} > 5$.

The efficiencies are parametrised with respect to various event properties, including the lepton momentum, its angular distance to the closest jet and the jet and b-tag multiplicity of the event. Systematic uncertainties of the prompt lepton fake estimate are related to the application of the tag-and-probe method, the normalisation uncertainties of W/Z+jets contributions and to the parametrisation of the efficiencies. Differences of fake and non-prompt lepton compositions in the event sample used to measure the fake efficiencies on the one hand, and the event sample of the present analysis on the other hand, are taken into account as well. Overall the systematic uncertainty amounts to 50% at most. Figure 3.3 presents distributions of m_{T}^{W} in event samples relevant to the present measurement. Indeed, the fakes of prompt leptons are modelled properly.

3.1.3. Event selection

Measured as well as simulated events are selected with regard to the desired measurement by means of kinematic cuts. The aim of this selection is to increase the fraction of signal events contained in the event sample under study, which in turn increases the sensitivity of the measurement to the signal process. Still, the signal fraction turns out to be relatively small, which makes the use of a more powerful discrimination approach necessary. This is the method of total event likelihoods to be discussed in the next section. However, the usage of kinematic cuts prior to this advanced discrimination is necessary in order to reduce the absolute number of background events.

As outlined above, the s-channel single top-quark events to be identified in the context of the measurement consist of two b-tagged jets and a W boson decay which results in a charged lepton and missing transverse momentum caused by a neutrino. In general, large transverse momenta are involved. These properties are shared by all of the selected events. Furthermore, three different regions of event kinematics are defined. The signal region provides sensitivity to s-channel single top-quark production, while two other regions are used for the purpose of validating the modelling of the dominant W+jets and $t\bar{t}$ backgrounds, respectively. The applied kinematic cuts are discussed in the following. Tab. 3.1 summarises them. Most of the cuts which define the three kinematic regions have already been used in a previous analysis based on the same data set, where Boosted Decision Trees were used to search for s-channel single top-quark production [116, 169].

General requirements

- **Primary vertex:** Every selected event is required to contain at least one primary vertex which involves at least five tracks.
- **Rejection of calorimeter noise events:** Events are rejected if they contain a jet with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ which is associated with noisy calorimeter cells.
- **One electron or muon, $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $|\eta| < 2.5$:** The presence of exactly one electron or muon is required. The reconstruction of τ leptons is relatively inefficient compared to the electron and muon reconstruction. Therefore τ leptons are not reconstructed explicitly, but leptonic τ decays can contribute to the selection. Electron candidates are selected with regard to their transverse energy E_{T}, defined by means of the energy of their electromagnetic energy cluster in the calorimeter and the pseudorapidity of their track, $E_{\mathrm{T}} = E_{\mathrm{Cluster}}/\cosh\eta_{\mathrm{track}} > 30\,\mathrm{GeV}$. This value is greater than the trigger threshold of the relevant `EF_e24vhi_medium1` item. The requirement of an electron track restricts the selection to $|\eta| < 2.5$. The transition region between the EM calorimeter barrel and endcaps, where the instrumentation is reduced, $1.37 < |\eta_{\mathrm{Cluster}}| < 1.52$, is excluded from the selection. Muon candidates need to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ for a combined inner detector and muon spectrometer track. Again, this value is above the threshold of the relevant trigger item, which is `EF_mu24i_tight` in this case. As in the electron case, the requirement of an inner detector track restricts the selection to $|\eta| < 2.5$.
- **Anti-k_{T} jets, $R = 0.4$:** Jets are reconstructed using the anti-k_{T} algorithm with a distance parameter of $R = 0.4$. Considered jets need to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$. In order to make use of b-tagging, inner tracking detector information is needed, which restricts the jet selection to $|\eta| < 2.5$. In order to reject jets caused by additional pile-up collisions, a jet vertex fraction of JVF > 0.5 is required in case of soft jets, $p_{\mathrm{T}} < 50\,\mathrm{GeV}$, in the central detector region, $|\eta| < 2.4$, where there is appropriate tracking information available.
- **Multijet rejection, $E_{\mathrm{T}}^{\mathrm{miss}} > 35\,\mathrm{GeV}$, $m_{\mathrm{T}}^{W} > 30\,\mathrm{GeV}$:** In order to reduce the amount of background caused by fakes of prompt leptons, also called multijet background, a missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}} > 35\,\mathrm{GeV}$ and a transverse W boson mass $m_{\mathrm{T}}^{W} > 30\,\mathrm{GeV}$ are required.

Signal region

- **Two b-tagged jets, $p_{\mathrm{T}}^{\mathrm{jet\,1}} > 40\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{jet\,2}} > 30\,\mathrm{GeV}$:** In order to match the generic signal signature, the signal region is defined by the requirement of exactly two jets. The leading jet needs to have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, while for the sub-leading jet $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ is required. The higher p_{T} threshold in case of the leading jet avoids sizeable modelling uncertainties present in lower jet p_{T} configurations. Both jets must be tagged as b-jets given the 70% efficiency working point of the MV1 algorithm. The presence of additional jets with $25\,\mathrm{GeV} < p_{\mathrm{T}} < 30\,\mathrm{GeV}$ is not allowed, as higher jet multiplicities are a characteristic feature of the $t\bar{t}$ background. Furthermore, events are rejected if they contain a jet which has $|\eta| > 2.5$. This cut reduces the amount of the t-channel single top-quark background, where forward scattering occurs frequently.

- **Di-lepton veto:** Further action is taken to reduce the amount of the dominant $t\bar{t}$ background. This process can contribute via single lepton events where a hadronic W boson decay is missed or via di-lepton events where a leptonic W boson decay is missed. A dedicated study reveals that about 75% of the $t\bar{t}$ events contained in the signal region described so far are di-lepton events [170]. Corresponding results are presented in App. C.2. In these events, one of the charged leptons resulting from a W boson decay does not meet the tight selection requirements. In order to reduce the amount of this background contribution, loosened electron and muon selection requirements are employed. They are given by the definitions of loose electrons and muons used in the context of the prompt lepton fake estimation discussed in Sec. 3.1.2. Furthermore, a threshold of $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ is used. Events which contain one of these loose leptons in addition to the one fulfilling the tight selection requirements are rejected. In consequence, the $t\bar{t}$ background is reduced by a considerable amount, while the number of signal events is essentially unchanged. This cut considerably improves the event selection compared to the previous analysis performed by the ATLAS collaboration [116, 169].

W+jets validation region

- **Two b-tagged jets, $p_{\mathrm{T}}^{\mathrm{jet\,1}} > 40\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{jet\,2}} > 30\,\mathrm{GeV}$, loosened b-tagging:** Another selection serves the purpose of validating the modelling of background processes, in particular of W+jets production. The selection is identical with the

Cuts \ Regions	Signal	$t\bar{t}$ validation	W+jets validation
Lepton	$N_e = 1$ or $N_\mu = 1$, $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$		
Di-lepton veto	no 2nd loose e or μ, $p_{\mathrm{T}} > 5\,\mathrm{GeV}$	none	
Jets	$N = 2$, $p_{\mathrm{T}}^{\mathrm{jet}\,1} > 40\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{jet}\,2} > 30\,\mathrm{GeV}$	$N = 4$, $p_{\mathrm{T}} > 25\,\mathrm{GeV}$	$N = 2$, $p_{\mathrm{T}}^{\mathrm{jet}\,1} > 40\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{jet}\,2} > 30\,\mathrm{GeV}$
b-tags	$N = 2$, MV1@70%		$N = 2$, MV1@80%, signal region excluded
Multijet rejection	$E_{\mathrm{T}}^{\mathrm{miss}} > 35\,\mathrm{GeV}$, $m_{\mathrm{T}}^{W} > 30\,\mathrm{GeV}$		

Table 3.1.: Event selection of the signal and validation regions. "MV1" denotes the employed b-tagging algorithm, where different selection efficiencies are distinguished.

one of the signal region except for a relaxation of the b-tagging requirement. The two selected jets need to be tagged as b-jets given the 80% efficiency working point of the MV1 algorithm, while the signal region is excluded from this selection, as illustrated in Fig. 3.4.

$t\bar{t}$ validation region

- **Four jets, $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, $|\eta| < 2.5$, two b-tags:** Top-quark pair production is the dominant background in the search for s-channel single top-quark production. For this reason, a dedicated validation region is defined in order to investigate the modelling of this process. In accordance with the generic signature of single lepton $t\bar{t}$ events, the presence of exactly four jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$ is required. Two of the jets must be b-tagged given the MV1 70% efficiency working point. Apart from these requirements, the selection is identical with the one of the signal region.

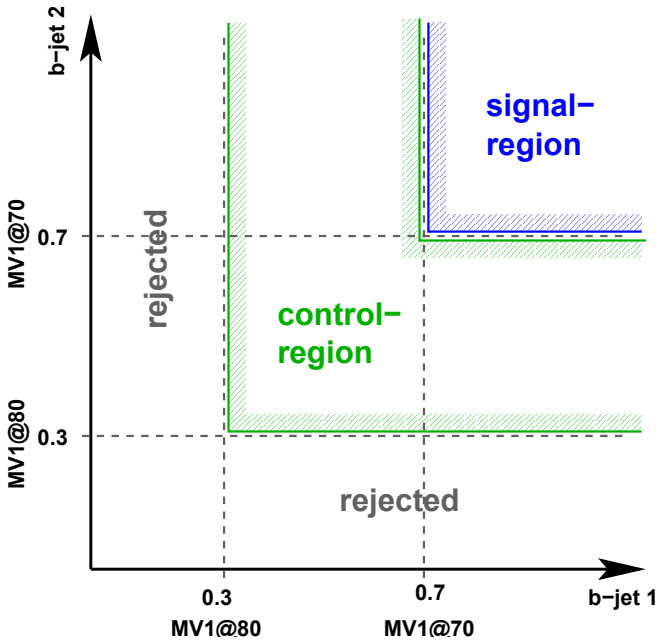


Figure 3.4.: Definition of the W+jets validation region. A relaxed b-tagging requirement is used in order to increase the fraction of W+jets events. The signal region is excluded from the selection.

Selected event display

In order to illustrate the collision events under study, Fig. 3.5 shows an example of a selected event in the signal region. The leading jet comes with a transverse momentum of $p_{\mathrm{T}} = 410\,\mathrm{GeV}$. This large momentum is partially balanced by another jet and a muon which emerge on the opposite side of the detector and have transverse momenta of $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ and $p_{\mathrm{T}} = 60\,\mathrm{GeV}$, respectively. The remaining imbalance is identified as missing transverse momentum, which amounts to $E_{\mathrm{T}}^{\mathrm{miss}} = 230\,\mathrm{GeV}$. The system of the muon and the missing transverse momentum is a W boson decay candidate. Together with the neighbouring jet it forms a top-quark decay candidate. The harder jet, on the other hand, is likely to be induced by a b-quark produced in association with the top-quark. Hence this is a typical s-channel single top-quark candidate event. In fact, the TELL method, which is used to identify these events by means of an approximate signal probability (Sec. 3.2), assigns a particularly high discriminant value to this event, namely $P(S|X) = 0.89$.

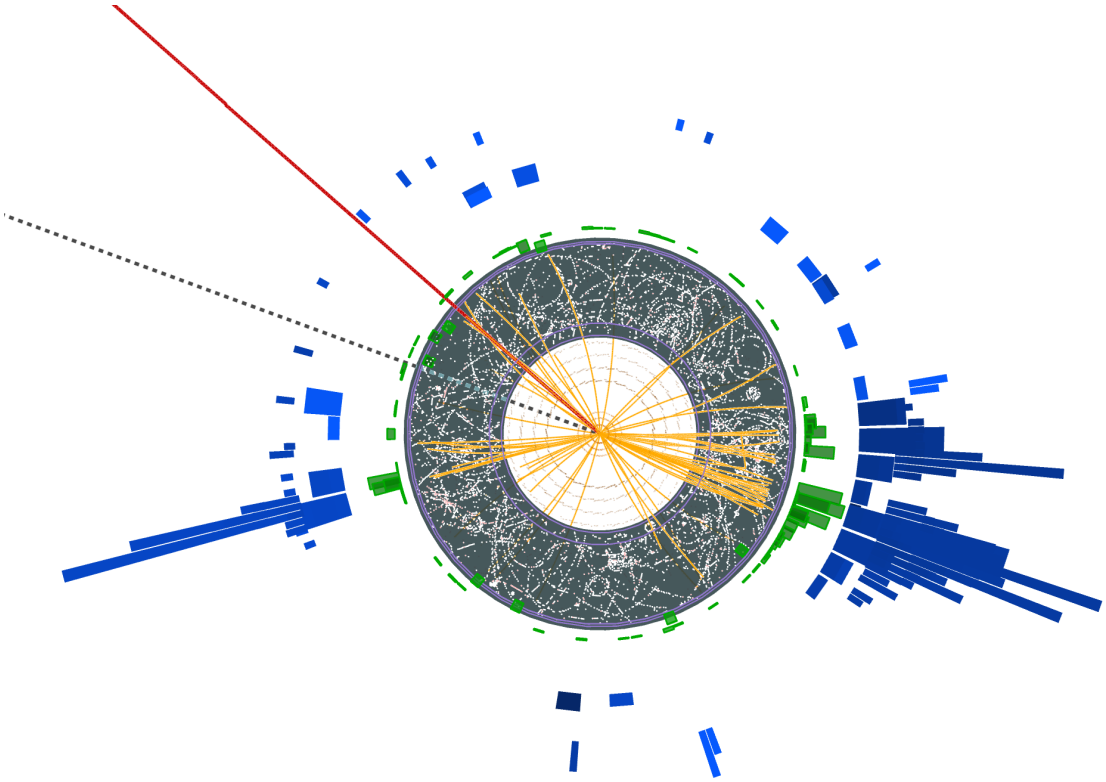


Figure 3.5.: Selected event display. The event is projected onto the transverse plane. The orange lines in the inner detector show tracks of charged particles. The red line shows a combined track, which is a muon candidate. The dashed line represents missing transverse momentum. The blue blocks show energy deposited in the calorimeter, while the green blocks represent reconstructed clusters of such deposits. A jet with a high transverse momentum shows up on the right-hand side. The other jet, the muon and the missing transverse momentum emerge on the other side of the detector. They form a top-quark decay candidate. Together with the leading jet, this is a typical signature of s-channel single top-quark production.

3.1.4. Event yields and control distributions

Before the desired measurement can be performed, the modelling of the data by the employed simulations and data-driven background estimates needs to be investigated. Since the measurement makes use of complete electron, muon and jet momenta in the context of the method of total event likelihoods, a good modelling of the corresponding momentum components must be achieved.

In order to determine the goodness of the modelling of the data, dedicated maximum likelihood fits are performed in all of the three regions discussed in the previous section. In each region, the distribution of the transverse W boson mass m_{T}^{W} is used to fit the model to the data. This variable is chosen because it is distributed differently for the two most abundant processes, namely W+jets production and top-quark pair production. The latter process involves di-leptonic events which result in larger missing transverse momentum values. This translates into larger values of m_{T}^{W}. Thus these fits allow one to constrain the individual normalisations of the two main scattering processes.

The likelihood function is based on a statistical model presented in detail in Sec. 3.4. It involves a scale factor for each process, which defines its normalisation. Regarding systematic uncertainties, these fits only involve the theoretical normalisation uncertainty of each simulated process, the normalisation uncertainty of the prompt lepton fake estimate, the statistical uncertainties of the predictions and the uncertainty of the integrated luminosity. A detailed discussion of these uncertainties is given below in Sec. 3.3. Other uncertainties are neglected here, because they are not required for the investigation of the goodness of the modelling, which is the only purpose of these fits. Within this context, the s-channel signal process is subject to a normalisation constraint according to the SM prediction in the same way as it is applied in case of the background processes. The desired s-channel measurement presented in Sec. 3.4 is independent of these fits.

The fit results are presented below. Tables 3.2, 3.3 and 3.4 provide the event yields in case of the signal, W+jets and $t\bar{t}$ region, respectively. Process normalisation scale factors obtained from the fits are applied in each case. The quoted uncertainties only involve statistical uncertainties. These are the statistical uncertainty of each simulation and data-driven estimate and the Poissonian error on the expected event yield added in quadrature. The W+jets prediction consists of light and heavy flavour contributions. The latter refers to events with b and c-quarks in the final state. In the context of the

Process	e	μ	ℓ
s-channel	250 ± 20	320 ± 20	570 ± 20
t-channel	550 ± 20	670 ± 30	1210 ± 40
Wt	180 ± 20	230 ± 20	410 ± 30
$t\bar{t}$	3600 ± 60	4580 ± 70	7970 ± 90
W+jets	2000 ± 70	2500 ± 100	4700 ± 200
Z+jets & di-boson	170 ± 20	250 ± 20	470 ± 30
Multi-jet	230 ± 20	220 ± 20	370 ± 20
Total expectation	6980 ± 100	8720 ± 160	15740 ± 200
Data	6914	8642	15556
S/B [%]	3.8	3.8	3.8

Table 3.2.: Event yields in the signal region. The yields are scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. In this context, all processes are constrained by their SM prediction within corresponding uncertainties, including the s-channel signal. Only statistical uncertainties are provided. The bottom row shows the signal-to-background ratios, S/B.

final measurement, all W+jets contributions are merged as the light flavour fraction is small. The simulations predict it to amount to approximately 9%, 30% and 4% of all W+jets events in the signal, W+jets and $t\bar{t}$ region, respectively. Accordingly, in the context of the control fits the W+jets merging is also done in case of the signal and $t\bar{t}$ region, but not in case of the W+jets region where the light flavour fraction is relatively large.

Figures 3.6 to 3.8 provide control distributions of several kinematic variables in all of the three regions. The electron and muon event selections are merged. The normalisation of each of the processes is scaled according to the corresponding fit result. Details on the fit results as well as further control distributions are provided in App. C. Overall a good modelling of the data is achieved in every case. This observation creates confidence in the validity of the employed modelling of scattering processes.

Process	e	μ	ℓ
s-channel	78 ± 9	100 ± 10	180 ± 10
t-channel	380 ± 20	470 ± 20	850 ± 30
Wt	180 ± 20	210 ± 20	390 ± 30
$t\bar{t}$	2170 ± 50	2530 ± 50	4620 ± 70
W+light jets	710 ± 80	900 ± 90	1600 ± 120
W+heavy flavour	2750 ± 80	3810 ± 100	6870 ± 130
Z+jets & di-boson	280 ± 20	270 ± 20	500 ± 30
Multi-jet	270 ± 20	320 ± 20	490 ± 30
Total expectation	6810 ± 130	8600 ± 150	15490 ± 200
Data	6881	8577	15458
S/B [%]	1.2	1.2	1.2

Table 3.3.: Event yields in the W+jets region. The yields are scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. In this context, all processes are constrained by their SM prediction within corresponding uncertainties, including the s-channel signal. Only statistical uncertainties are provided. The bottom row shows the signal-to-background ratios, S/B.

Process	e	μ	ℓ
s-channel	39 ± 6	50 ± 7	89 ± 9
t-channel	310 ± 20	400 ± 20	710 ± 30
Wt	560 ± 30	690 ± 30	1250 ± 50
$t\bar{t}$	17800 ± 100	22400 ± 200	40100 ± 200
W+jets	1290 ± 50	840 ± 40	2250 ± 70
Z+jets & di-boson	140 ± 10	100 ± 10	250 ± 20
Multi-jet	60 ± 10	50 ± 10	110 ± 20
Total expectation	20160 ± 150	24530 ± 170	44750 ± 230
Data	20328	24776	45104
S/B [%]	0.2	0.2	0.2

Table 3.4.: Event yields in the $t\bar{t}$ region. The yields are scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. In this context, all processes are constrained by their SM prediction within corresponding uncertainties, including the s-channel signal. Only statistical uncertainties are provided. The bottom row shows the signal-to-background ratios, S/B.

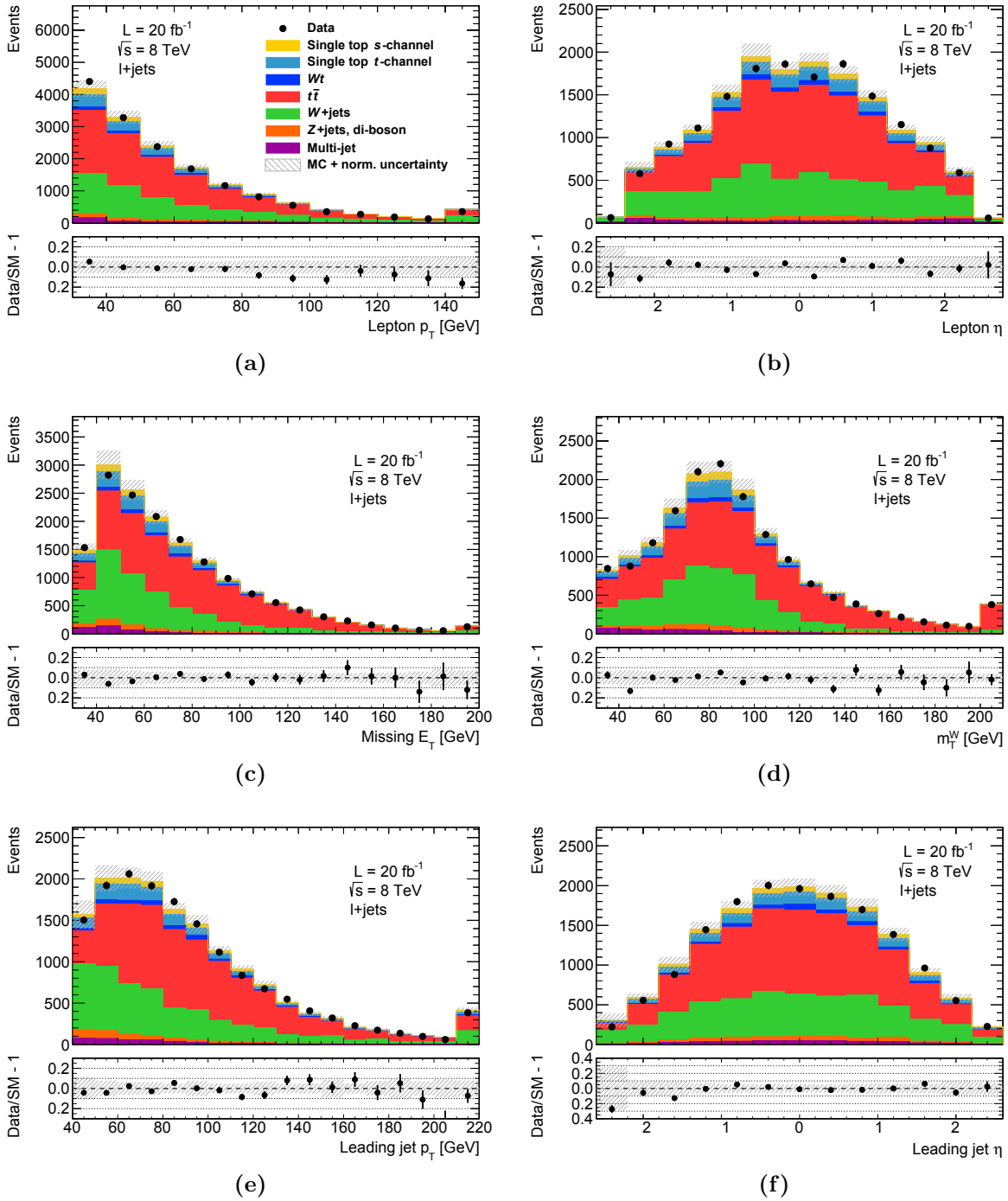


Figure 3.6.: Kinematic distributions in the signal region. Each process is scaled according to the result of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands involve the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.

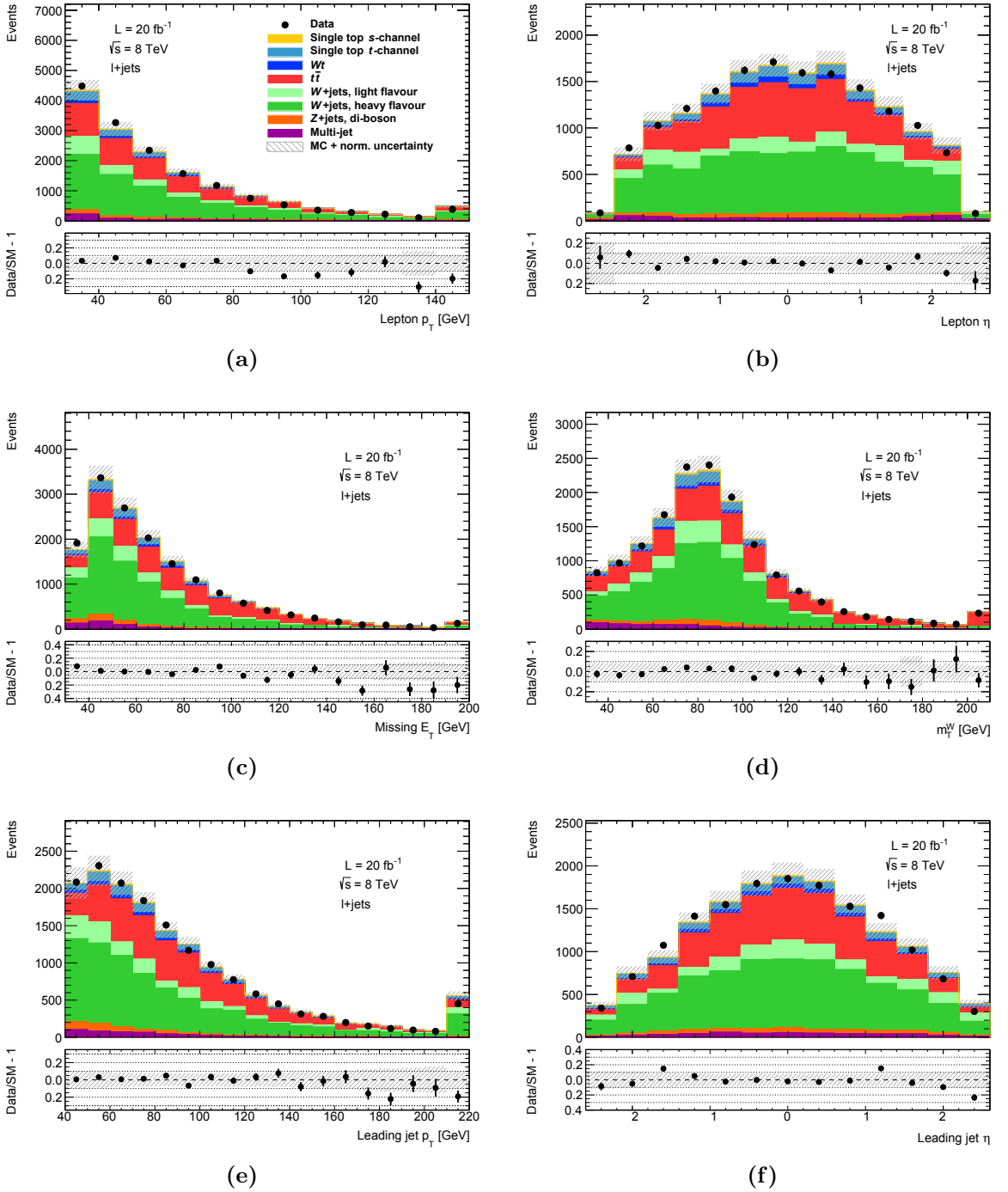


Figure 3.7.: Kinematic distributions in the W+jets region. Each process is scaled according to the result of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands involve the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.

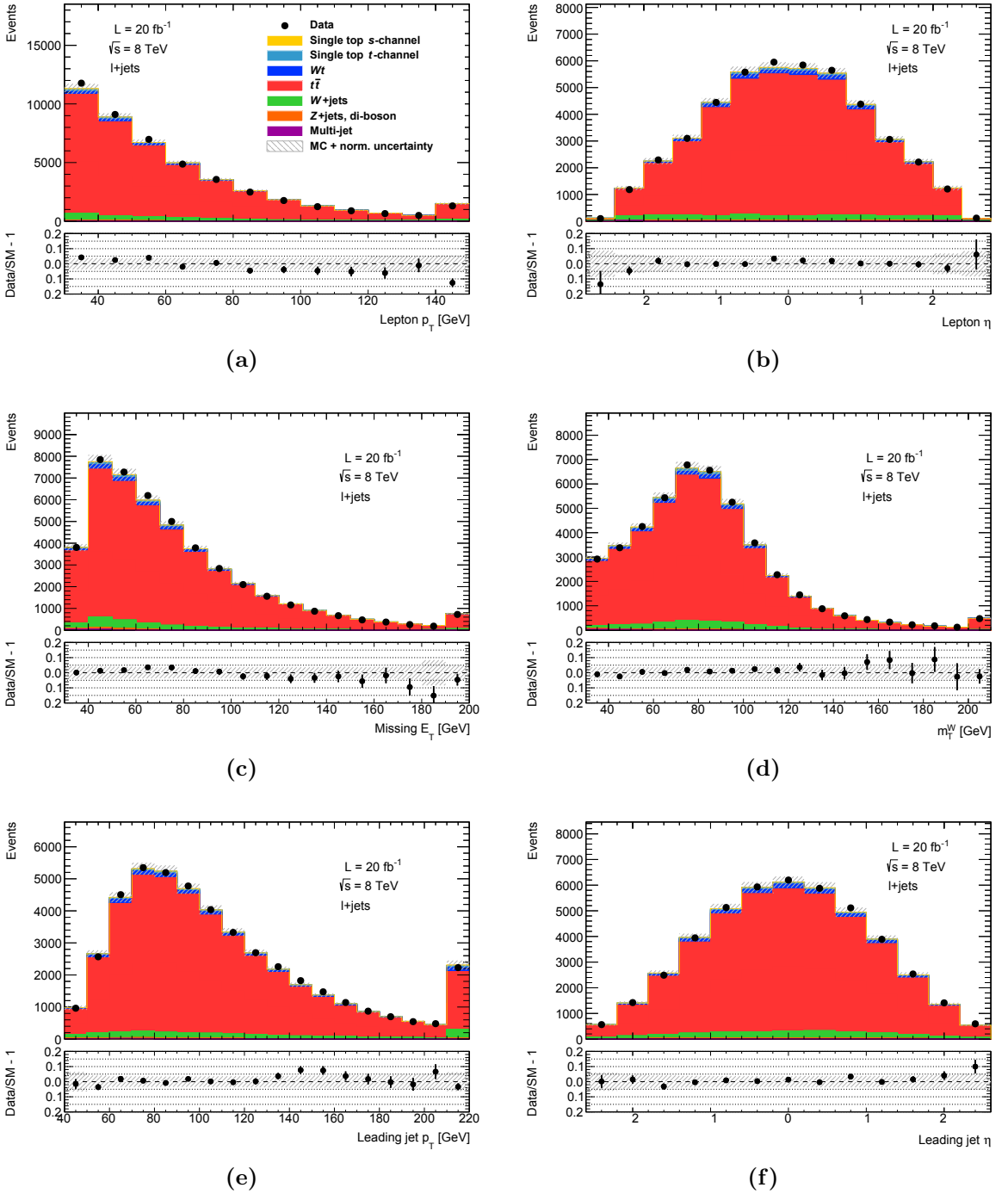


Figure 3.8.: Kinematic distributions in the $t\bar{t}$ region. Each process is scaled according to the result of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands involve the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.

3.2. The method of total event likelihoods

In order to discover a scattering process of interest, the number of corresponding collision events under study must be sufficiently large compared to the uncertainty of the number of background events. The kinematic cuts discussed in the previous section define the absolute number of signal events within the considered set of collisions. Given the large number of background events and the involved uncertainties, the mere kinematic selection is not powerful enough to allow for the identification of a possible signal. Hence a procedure is needed which separates the signal from its backgrounds.

The most general approach to this challenge is to make use of the probability density function (Pdf) of collision events, given a hypothesis of a scattering process. In this context, events are defined by the four momenta of jets, charged leptons and missing transverse momentum. In principle one could generate the Pdf by means of MC event generators. In fact, this approach would provide the maximum significance in any case, as outlined below. However, in general the Pdf involves several momentum dimensions. For this reason, the optimal approach cannot be taken in practice, given limited amounts of computing power.

In order to reduce the computing demand, approximations can be used. This is the basic idea of the "Dynamical Likelihood Method", or method of total event likelihoods (TELL), originally proposed in 1988 [171, 172]. First of all, the scattering process is factorized in a way which is similar to the factorization of MC event generation discussed in Sec. 1.4. Concerning the hard scattering process, the description is restricted to the leading order of perturbation theory, as far as current applications are concerned. Furthermore, phenomena which involve lower energy scales like parton showers, hadronization and the detector response are summarised and approximated by means of parametrisations, called transfer functions. They depend on the measured momenta. Finally, the Pdf is given by a convolution of the differential cross section of the hard scattering process with the transfer functions.

In principle this method can be applied to measurements of any scattering process which can be described by means of perturbation theory. Furthermore, it can be used for measurements of various observables. On the one hand, it allows for cross section measurements by means of a separation between signal and background processes. This is how the method has been used in single top measurements at the Tevatron. Combined with other separation methods, namely machine learning techniques, it allowed for the

first observation of single top-quark production [73, 74]. The same is true for the evidence of s-channel single top-quark production found by the D0 collaboration [113], discussed in Sec. 1.3.4. On the other hand, model parameters can be measured by means of this method as well. The most famous example is given by measurements of the top-quark mass m_t at the Tevatron. In fact, the most precise single measurement of m_t to date is based on this method [173]. Another example is given by measurements of W boson polarisations in top-quark decays performed by CDF [174].

However, the method requires considerable amounts of computing power. In consequence, there are only a few cases in which it has been applied by the ATLAS and CMS collaborations so far. In particular at the beginning of collision data taking, the work of the two collaborations was characterized by a high pace. Meanwhile the need to extract more information from the available data sets arose, so that the use of this method is motivated. Examples are given by searches for the associated production of a pair of top-quarks and a Higgs boson [175, 176].

A remark must be made concerning the name of the method. In accordance with the aim to approximate a Pdf, the original proposal introduced the name "Dynamical Likelihood Method" [171, 172]. Subsequently, the name "Matrix Element Method" has been established instead for obviously no reason. This name expresses nothing but the fact that Quantum Mechanics, or a matrix element is involved. Thus it is trivial and lacks any significant meaning. Clearly, a meaningful name following the original proposal is desired. At the same time, confusion should be avoided regarding the popular "Kinematic Likelihood Fitter" [177], a tool to perform kinematic fits which are based on selected event variables rather than the total event information. In this thesis the method of choice is called the method of "*T*otal *E*vent *L*ike*l*ihoods" (TELL), because that is what it is about.

In the present section, the theoretical motivation of TELLs is discussed in more detail first, followed by a description of the involved approximations. Subsequently, a concrete computer implementation of the general concept is presented. The employed ATLAS transfer functions are provided next. This is followed by a comparison of the event likelihoods obtained in experimental data and corresponding predictions relevant to the desired measurement. Finally, the discrimination of signal and background events is presented. In addition, a closer look is taken at the efficiency of the likelihood computations. Also the quality of the likelihoods is estimated.

3.2.1. Ansatz

Neyman–Pearson Lemma

The aim of the TELL method in the context of this search is to map the momenta of reconstructed objects from a collision event to a discriminant variable which allows one to distinguish different scattering processes. While many other means to achieve this goal have been developed, in particular machine learning techniques, it has never been proven that one of them generally outperforms all others. Two issues need to be considered when it comes to signal discrimination. First, there is the degree of separation between signal and background processes. The higher the separation, the smaller the impact of statistical uncertainties. Secondly, there is the impact of systematic uncertainties on the discriminant. Concerning the separation, there is a clear answer to the question which quantity is optimal in order to distinguish two simple hypotheses like "there is a signal" or "there is no signal". It is given by the Neyman–Pearson lemma [178]. Let $H_0 : \mu = 0$ (background) and $H_1 : \mu = 1$ (signal) be simple hypotheses, where μ denotes the signal event frequency with respect to a reference model. The outcome $\mathbf{X}$ of an experiment, namely a set of momenta measured in collision events, is mapped onto a test statistic $Q(\mathbf{X}) \in \mathbb{R}$. A rejection region $R_{\not{H_0}} \subset \mathbb{R}$ is defined in order to determine which hypothesis needs to be chosen. The test rejects H_0 in favour of H_1 if and only if $Q(\mathbf{X}) \in R_{\not{H_0}}$. In other words, the signal is assumed to exist in this case. Otherwise the null-hypothesis H_0 is accepted. The Neyman–Pearson test involves a fixed probability to reject H_0 if it is true,

$$\alpha = P\left(Q(\mathbf{X}) \in R_{\not{H_0}} | H_0\right) \ . \tag{3.7}$$

It is called the significance level of the test and constitutes a boundary condition. An optimal choice of the rejection region $R_{\not{H_0}}$ can be defined as the one which gives the highest probability to reject H_0 if it is wrong. Since in this case H_1 is accepted, the power of a hypothesis test is defined as

$$\beta = P\left(Q(\mathbf{X}) \in R_{\not{H_0}} | H_1\right) \ . \tag{3.8}$$

Accordingly, optimising a hypothesis test means to maximize the power β given a significance level α. In other words, for a given type I error probability α the type II error probability $1 - \beta$ must be minimal. The statement of the Neyman–Pearson lemma

is the fact that the optimal test statistic is given by the ratio of probability densities of $\mathbf{X}$, also called the likelihood ratio. H_0 is rejected in favour of H_1 if and only if this ratio is larger than a constant value k_α, which defines the significance level α of the test,

$$Q(\mathbf{X}) = \frac{P(\mathbf{X}|H_1)}{P(\mathbf{X}|H_0)} , \tag{3.9}$$

$$R_{\not H_0} = \{Q(\mathbf{X}) \,:\, Q(\mathbf{X}) > k_\alpha\} \ . \tag{3.10}$$

Application of the Neyman–Pearson Lemma

In order to relate the Neyman–Pearson lemma to a search for a scattering process, let $\mathbb{X} = \{\mathbb{X}_i\}_{i=1}^n$ be the set of all measured events in a given data set. Here an event $\mathbb{X}_i$ is defined by a set of measured momenta and the scattering process which it originates from,

$$\mathbb{X}_i = (\{\text{measured momenta } \mathbf{p}_\text{i}\}, \text{process } H_i) \ . \tag{3.11}$$

The sample space of events is the space of all possible outcomes of a collision $\Omega = \{\mathbb{X}\}$. The set of all scattering processes is divided into signal and background processes,

$$\bigcup_i \{H_i\} = \{S_1, S_2, \ldots, S_{N_S}, B_1, B_2, \ldots, B_{N_B}\} \ . \tag{3.12}$$

It is assumed that no interference occurs between these processes. The definition of several signal contributions S_i allows one to split the overall signal process into different channels, e.g. with regard to measured lepton flavours and charges or the number of final state jets.

Given an event, only the momenta $\mathbf{p}$ can be measured, while it is in general unknown which of the scattering processes H has occured. What can be computed are the probability densities of measurements of momenta $X = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n)$, which are called "event likelihood" in the following,

$$\mathcal{P}(X|H) \,:\, \text{Pdf of the momentum measurement } X \text{ given the process } H \tag{3.13}$$

In other words, $\int_\text{M} \mathcal{P}(X|H)\text{d}X = \int_\text{M} \mathcal{P}\left(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n|H\right) \text{d}^n\mathbf{p}$ is the probability to measure momenta in a detector region M given the scattering process H. The momenta $\{\mathbf{p}\}$ might be restricted to a certain detector region $D \subset \{X\}$. This affects the normalisation

of the event likelihood,

$$\int_D \mathcal{P}(X|H)\,\mathrm{d}X = 1\ . \tag{3.14}$$

Given these entities, the Neyman–Pearson test statistic can be constructed. The null-hypothesis H_0 is defined by the absence of the signal processes. On the other hand, H_1 is the signal hypothesis,

$$H_0: \qquad \mathcal{P}_0(X) \;=\; \frac{1}{\sum_i B_i}\sum_i B_i\,\mathcal{P}(X|B_i)\ , \tag{3.15}$$

$$H_1: \qquad \mathcal{P}_1(X) \;=\; \frac{1}{\sum_i S_i + \sum_i B_i}\left(\sum_i S_i\,\mathcal{P}(X|S_i) + \sum_j B_i\,\mathcal{P}(X|B_i)\right)\ . \tag{3.16}$$

S_i and B_i denote the mean number of expected signal and background events, respectively. Furthermore, let $S = \sum_i S_i$ and $B = \sum_i B_i$ denote the total number of expected signal and background events, respectively. The measured total number of events is denoted by n. Given these definitions, the likelihood ratio test statistic is [179],

$$\begin{aligned} Q(\mathbf{X}) = \frac{P(\mathbf{X}|H_1)}{P(\mathbf{X}|H_0)} &= \frac{\mathrm{Pois}(n|S+B)}{\mathrm{Pois}(n|B)}\frac{\prod_{k=1}^n \mathcal{P}_1(X_k)}{\prod_{k=1}^n \mathcal{P}_0(X_k)} = e^{-S}\left(\frac{S+B}{B}\right)^n \prod_{k=1}^n \frac{\mathcal{P}_1(X_k)}{\mathcal{P}_0(X_k)} \\ &= e^{-S}\prod_{k=1}^n\left(1 + \frac{\sum_i S_i\,\mathcal{P}(X_k|S_i)}{\sum_j B_j\,\mathcal{P}(X_k|B_j)}\right)\ . \end{aligned} \tag{3.17}$$

Here $\mathrm{Pois}(n|\mu) = \mu^n e^{-\mu}/n!$ denotes the Poisson distribution. In view of applications discussed below, this expression is translated into

$$\begin{aligned} q(\mathbf{X}) =\ & \ln\left(Q(\mathbf{X})\right) = -S + \sum_{k=1}^n \ln\left(1 + \frac{\sum_i S_i\,\mathcal{P}(X_k|S_i)}{\sum_j B_j\,\mathcal{P}(X_k|B_j)}\right) \\ =\ & -S - \sum_{k=1}^n \ln\left(\frac{\sum_j B_j\,\mathcal{P}(X_k|B_j)}{\sum_i S_i\,\mathcal{P}(X_k|S_i) + \sum_j B_j\,\mathcal{P}(X_k|B_j)}\right) \\ =\ & -S - \sum_{k=1}^n \ln\left(1 - \frac{\sum_i S_i\,\mathcal{P}(X_k|S_i)}{\sum_i S_i\,\mathcal{P}(X_k|\ S_i) + \sum_j B_j\,\mathcal{P}(X_k|B_j)}\right)\ . \end{aligned} \tag{3.18}$$

There is a simple interpretation to the likelihood ratio which appears at the end of Eq. 3.18. Let $P(H)$ be the a-priori probability of the process H,

$$\begin{aligned} P(S_i) &= \frac{S_i}{\sum_j S_j + \sum_j B_j} \ , \\ P(B_i) &= \frac{B_i}{\sum_j S_j + \sum_j B_j} \ . \end{aligned} \tag{3.19}$$

With the help of these quantities, the interpretation of the likelihood ratio becomes clear,

$$q(\mathbf{X}) = -S - \sum_{k=1}^{n} \ln\left(1 - P(S|X_k)\right) \ , \tag{3.20}$$

$$P(S|X_k) = \frac{\sum_i P(S_i)\,\mathcal{P}(X_k|S_i)}{\sum_i P(S_i)\,\mathcal{P}(X_k|S_i) + \sum_j P(B_j)\,\mathcal{P}(X_k|B_j)} \ . \tag{3.21}$$

With $\mathcal{P}(X|H)$ being the probability density for a measurement of momenta X in an event given the scattering process H and $P(H)$ being the a-priori probability of the latter, $P(S|X)$ is the probability of the signal process to occur, given the measured event X. This interpretation follows from Bayes' theorem. In summary, $P(S|X)$ provides all the information necessary for an optimal hypothesis test.

Approximation of the event likelihood

In principle the function defined by Eq. 3.21 can be determined by means of MC simulations. Concerning the a-priori probabilities $P(H_i)$ this is simple. They are given by the relative frequencies of events of the process H_i as predicted by the simulations. On the contrary, the likelihoods $\mathcal{P}(X|H)$ cannot be determined by MC simulations in practice, due to the high computing demands of this task. This is the reason why there are many different means to search for scattering processes, as presented in Tab. 3.5. In the most simple cut and count approach, $\mathcal{P}(X|H)$ is integrated over an appropriate region of phase space. More flexibility is gained by marginalisations of $\mathcal{P}(X|H)$, where the integration is restricted to a subset of the phase space, which leads to a kinematic distribution. In challenging scenarios, machine learning techniques are used in order to define the residual variable of the Pdf. The most general applicable approach is the TELL method. It holds on to the usage of the Pdf $\mathcal{P}(X|H)$ by approximating it.

The approximation of the TELL method constitutes a factorization. One part of this factorization is the hard scattering process, described by the leading order fully

Type	Discriminant	Usage
Integration	$N(\Omega) = \int_\Omega \mathrm{d}X\ \mathcal{P}(X\|H)$	Choice of a phase space region $\Omega \subset \{X\}$, *cut and count*
Marginalization	$\frac{\mathrm{d}N}{\mathrm{d}y} = \int \mathrm{d}X \mathcal{P}(X\|H)\delta(y - g(X))$	Choice of a function $g(X)$, e.g. based on *machine learning*
Approximation	$\mathcal{P}(X\|H) \approx \int \mathrm{d}\Phi \frac{1}{\sigma_H}\frac{\mathrm{d}\sigma_H}{\mathrm{d}\Phi} T(X\|\Phi)$	Approximation of the probability density $\mathcal{P}(X\|H)$ - *TELL*

Table 3.5.: Overview of signal separation methods. $\mathcal{P}(X|H)$ denotes the probability density of collision events X given the scattering process H. It is used in multiple ways in order to search for scattering processes of interest. The TELL method, chosen in the present analysis, is the most general applicable approach. It makes use of complete collision events rather than selected kinematic variables.

differential cross section $\mathrm{d}\sigma_H/\mathrm{d}\Phi$. All subsequent effects including strong interactions at lower scales and the detector response are modelled by parametrisations called transfer functions $T(X|\Phi)$. In order to approximate $\mathcal{P}(X|H)$, the differential cross section of the hard scattering is convoluted with the transfer functions,

$$\begin{aligned}\mathcal{P}(X|H) = & \int \mathrm{d}\Phi \frac{1}{\sigma_H}\frac{\mathrm{d}\sigma_H}{\mathrm{d}\Phi} T(X|\Phi) \\ = & \frac{1}{\sigma_H}\sum_{a,b}\int \mathrm{d}x_1\mathrm{d}x_2 \prod_i \frac{\mathrm{d}^3\mathbf{p}_i}{(2\pi)^3 2E_i}(2\pi)^4\delta^{(4)}\left(x_1P_1 + x_2P_2 - \sum_i p_i\right) \\ & \cdot f_a(x_1)f_b(x_2)\frac{1}{2x_1x_2s}\left|\mathcal{M}_H^{a,b}\right|^2 T_H\left(X|\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\right) \ . \end{aligned} \tag{3.22}$$

The normalized fully differential cross section $1/\sigma_H \cdot \mathrm{d}\sigma_H/\mathrm{d}\Phi$ gives the probability density of the scattering process H to lead to a final state $\Phi = (x_1, x_2, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n)$ in the leading order prescription. The transfer function $T(X|\Phi)$ gives the probability density of this final state to lead to a measured event X. In consequence, the integration over final states Φ gives $\mathcal{P}(X|H)$, the TELL of the measurement X.

Concerning the modelling of the detector response, the transfer function $T(X|\Phi)$ needs to describe three different effects, namely possible ambiguities concerning the assignment of parton level momenta to reconstructed momenta, finite resolutions and finite reconstruction efficiencies. First of all, the transfer function is factorized into single

particle contributions where it is assumed that each final state particle momentum $\mathbf{p}$ is either matched to a reconstructed momentum $\mathbf{p}^{\mathrm{rec}}$, or not matched at all. For every pair of matched objects, a resolution function $W_{\mathrm{res}}(\mathbf{p}^{\mathrm{rec}}|\mathbf{p})$ enters. In order to simplify the computation, it is assumed that directions are measured precisely, which leads to δ-functional resolutions of angles. On the contrary, the employed energy resolutions usually have finite widths. This choice is in accordance with typical reconstruction capabilities. Overall, the resolution function reads

$$W_{\mathrm{res}}(\mathbf{p}^{\mathrm{rec}}|\mathbf{p}) = \frac{1}{(E^{\mathrm{rec}})^2} W^E_{\mathrm{res}}(E^{\mathrm{rec}}|E) \cdot \delta\left(\cos\vartheta^{\mathrm{rec}} - \cos\vartheta\right) \cdot \delta\left(\varphi^{\mathrm{rec}} - \varphi\right) \ . \tag{3.23}$$

Here a factor $1/(E^{\mathrm{rec}})^2$ is introduced for the purpose of normalisation discussed below. Apart from resolutions, reconstruction efficiencies $\varepsilon(\mathbf{p})$ enter for every matched particle. Every unmatched particle, on the other hand, comes with the counter efficiency $1-\varepsilon(\mathbf{p})$. The normalisation of the transfer function is related to the detector acceptance region D which appears above in Eq. 3.14. In order to simplify the computations, no restrictions are imposed on D in the following. For instance, $D = \mathbb{R}^{3n \,\cdot\, N_p}$ in case of a measurement of N_p massless objects. Accordingly, the normalisation condition of the resolution functions reads

$$\begin{aligned} &\int_{\mathbb{R}^3} \mathrm{d}^3\mathbf{p}^{\mathrm{rec}}\, W_{\mathrm{res}}(\mathbf{p}^{\mathrm{rec}}|\mathbf{p}) = 1 \ , \\ \Leftrightarrow &\int_{\mathbb{R}} \mathrm{d}E^{\mathrm{rec}}\, W^E_{\mathrm{res}}(E^{\mathrm{rec}}|E) = 1 \ . \end{aligned} \tag{3.24}$$

In order to combine the single particle transfer functions into the overall function $T(X|\Phi)$, all possible particle assignments need to be taken into account. The resulting average involves a matching of partons to jets, electrons to reconstructed electrons etc.

$$T(X|\Phi) = \frac{\tau_H}{N_{\mathrm{perm}}} \sum_{i\in\{\text{permutations}\}} \prod_{j\in\{\text{matched}\}} W_{\mathrm{res}}\left(\mathbf{p}^{\mathrm{rec}}_{ij}|\mathbf{p}_{ij}\right) \varepsilon(\mathbf{p}_{ij}) \prod_{k\in\{\text{unmatched}\}} \left(1 - \varepsilon\left(\mathbf{p}_{ik}\right)\right) \tag{3.25}$$

The sum over permutations and the products over matched and unmatched particles depend on the topology of the event X and on the process H. The coefficient τ_H in Eq. 3.25 serves the purpose of normalisation. Due to the normalisation of the resolutions according to Eq. 3.24, the integral of the complete transfer function over all of the

reconstructed momenta involves only efficiencies,

$$\begin{aligned}\int_D \mathrm{d}X\, T(X|\Phi) &= \frac{\tau_H}{N_{\text{perm}}} \sum_i \prod_j \underbrace{\int_{\mathbb{R}^{3n}} \mathrm{d}^3\mathbf{p}^{\text{rec}}\, W_{\text{res}}\left(\mathbf{p}_{ij}^{\text{rec}}|\mathbf{p}_{ij}\right)}_{=1} \varepsilon(\mathbf{p}_{ij}) \prod_k \left(1-\varepsilon\left(\mathbf{p}_{ik}\right)\right) \\ &= \frac{\tau_H}{N_{\text{perm}}} \sum_i \prod_j \varepsilon(\mathbf{p}_{ij}) \prod_k \left(1-\varepsilon\left(\mathbf{p}_{ik}\right)\right) \\ &= \tau_H \varepsilon_H(\Phi)\ . \end{aligned} \tag{3.26}$$

The definition of ε_H leads to a compact expression for the likelihood normalisation constraint of Eq. 3.14,

$$\begin{aligned}\int_D \mathcal{P}(X|H)\,\mathrm{d}X = \underbrace{\frac{\tau_H}{\sigma_H}}_{=c_H} \int \mathrm{d}\Phi\, \frac{\mathrm{d}\sigma_H}{\mathrm{d}\Phi} \varepsilon_H(\Phi) \stackrel{!}{=} 1 \\ \Rightarrow c_H = \left(\int \mathrm{d}\Phi\, \frac{\mathrm{d}\sigma_H}{\mathrm{d}\Phi} \varepsilon_H(\Phi)\right)^{-1}\ . \end{aligned} \tag{3.27}$$

The constant coefficients σ_H and τ_H are combined into a single factor c_H which must be determined once for each process H according to this equation.

The integration of the event likelihood can be realized using Monte Carlo techniques. The implementation of the integrand is straightforward. Overall it must be noticed that this approach is subject to many approximations. The hard scattering process is only described in leading order. In particular, this means that the total transverse momentum of the overall system of particles is modelled poorly. One way to allow for finite transverse momenta of a system of particles of interest is to make use of higher order tree level diagrams. In this approach additional, unresolved particles need to be integrated. Recently, the extension of the method to the next-to-leading order of perturbation theory has been achieved [180]. The application of this concept in terms of an experimental analysis still needs to be addressed. Another shortcoming of the method is given by the simplicity of the parametrisation of strong interactions at low scales and of detector effects in terms of the transfer functions.

For these reasons, one cannot expect these event likelihoods to model measurements in a very detailed way. On the other hand, the method pays attention to all major aspects of collision events in a symmetric way. No particular variables are preferred but instead the whole set of measured momenta is taken into account including all of their correlations. At this point of the discussion the relative impact of the shortcomings and

advantages of the TELL method can hardly be estimated. Furthermore, the impact of systematic uncertainties on measurements cannot be discussed in general at this point. After all, the TELL method is well motivated by first principles. It involves considerable approximations, but handles measurements in a comprehensive way. The value of this approach remains to be assessed on the basis of its applications in the field.

3.2.2. Implementation

So far there are only a few tools which are capable of computing event likelihoods. Most notably, the MadWeight 5 code [181, 182] covers a wide range of scattering processes as it is interfaced with the scattering amplitudes provided by the MadGraph 5 code [127] and makes use of a general algorithm to integrate over phase spaces [183]. Disadvantages of MadWeight 5 are its lack of likelihood normalisations as well as reconstruction efficiencies regarding the transfer functions. Furthermore, this tool was not available when the present physics analysis was started. Therefore, a new tool which meets the requirements of this analysis has been developed in cooperation with the experimental particle physics research group at Humboldt-Universität zu Berlin [184]. It is dedicated to fast computations of total event likelihoods, which includes all of the building blocks introduced above. A particular advantage is given by its interface, which is both user-friendly and flexible. In addition, its modular code structure allows one to extend it easily if desired. The tool can be used to study various scattering processes relevant to single top-quark measurements and beyond. In this section, the elements of the likelihood computation implemented in the tool are discussed first, followed by a discussion of the resulting computer program. The following expression serves as a listing of the building blocks, summarising Eq. 3.22, 3.25 and 3.27,

$$\begin{aligned}\mathcal{P}(X|H) = & \int \mathrm{d}\Phi \, \frac{1}{\sigma_H} \frac{\mathrm{d}\sigma_H}{\mathrm{d}\Phi} T(X\,|\Phi) \\ = & \, c_H \frac{1}{N_{\text{perm}}} \sum_{i \in \{\text{permutations}\}} \int \mathrm{d}x_1 \mathrm{d}x_2 \prod_l \frac{\mathrm{d}^3\mathbf{p}_l}{(2\pi)^3 2E_l} (2\pi)^4 \delta^{(4)} \left(P_{\text{in}} - P_{\text{out}}\right) \\ & \cdot \sum_{a,\,b} f_a(x_1) f_b(x_2) \frac{1}{2x_1 x_2 s} \left| \mathcal{M}_H^{a,\,b} \right|^2 \\ & \cdot \prod_j W_{\text{res}} \left(\mathbf{p}_{ij}^{\text{rec}} | \mathbf{p}_{ij}\right) \varepsilon(\mathbf{p}_{ij}) \prod_k \left(1 - \varepsilon\left(\mathbf{p}_{ik}\right)\right) \ . \end{aligned} \tag{3.28}$$

Transfer function permutations

Before the likelihood computation can start, all possible matchings between the final state particles of the particular process on the one hand, and the physics objects of the measured event on the other hand must be identified. Charged leptons from W boson decays are matched to reconstructed electrons or muons of the same charge. Partons are matched to jets. By default, b-tagged jets are only assigned to b-quarks rather than light partons, because mis-tagging probabilities are small. This constraint can be removed if desired. There can also be particles which are not matched to a reconstructed object, e. g. two of the four partons of a single-lepton $t\bar{t}$ process in case of a two-jet event. As soon as all possible matchings, or permutations, are determined, the likelihood computation is performed successively for each permutation. The result is given as the sum of each of these contributions. The reason to integrate single permutations successively instead of having a single integration of all permutations at once is given by the fact that the peak structure of the integrand differs from one permutation to another. This is because the matching between particles and reconstructed objects affects the resolution functions, which are peak functions. Therefore the choice of integrating single permutations simplifies the Monte Carlo integration.

To illustrate the expression $T(X|\Phi)$, a concrete example relevant to the measurement presented in this thesis is shown here. Let s-channel single top-quark production with three outgoing partons be the scattering process H. Assuming a leptonic W boson decay, the final state denoted by Φ consists of a charged lepton, a neutrino, two b-quarks and a light parton radiated additionally, $q\bar{q}' \to \ell\nu b\bar{b}g$. The measured event X consists of a charged lepton, missing transverse momentum and two b-tagged jets. This configuration results in two permutations,

$$\begin{aligned}
\Phi &= \left(\mathbf{p}_\ell, \mathbf{p}_\nu, \mathbf{p}_b, \mathbf{p}_{\bar{b}}, \mathbf{p}_g\right) \ , \\
X &= \left(\mathbf{p}_{\ell,\mathrm{rec}}, E_\mathrm{T}^\mathrm{miss}, \mathbf{p}_{b\text{-jet}\,1}, \mathbf{p}_{b\text{-jet}\,2}\right) \ , \\
T(X|\Phi) &= W_\mathrm{res}^\ell(\mathbf{p}_{\ell,\mathrm{rec}}|\mathbf{p}_\ell)\varepsilon_\ell(\mathbf{p}_\ell)\varepsilon_{E_\mathrm{T}^\mathrm{miss}}(p_\mathrm{T}^\nu) \\
&\quad \cdot \left(1 - \varepsilon_\mathrm{jet}(\mathbf{p}_g)\right)\varepsilon_\mathrm{jet}(\mathbf{p}_b)\varepsilon_{b\text{-tag}}(\mathbf{p}_b)\varepsilon_\mathrm{jet}(\mathbf{p}_{\bar{b}})\varepsilon_{b\text{-tag}}(\mathbf{p}_{\bar{b}}) \\
&\quad \cdot \frac{1}{2}\left(W_\mathrm{res}^{b\text{-jet}}(\mathbf{p}_{b\text{-jet i}}|\mathbf{p}_b)W_\mathrm{res}^{b\text{-jet}}(\mathbf{p}_{b\text{-jet j}}|\mathbf{p}_{\bar{b}}) \ + \ [i \leftrightarrow j]\right) \ . \qquad (3.29)
\end{aligned}$$

Monte Carlo integration

The integration of likelihoods is performed by means of a Monte Carlo technique using the VEGAS algorithm [66, 67] as implemented in the CUBA library [185]. In general, the integral is evaluated multiple times. After each iteration, a grid defining preferred regions in random number space is refined. This refinement allows one to focus on regions in which the integrand is large, thereby accelerating the integration procedure. The number of integrand function calls per iteration and its increase from one iteration to another is set individually for each process. The algorithm keeps going until a target precision is reached or until a maximum number of integrand calls has been made. Integration parameters are optimised with regard to the required computation time and the precision of the resulting integral.

Phase space generation

The random numbers generated during the MC integration must be mapped onto the particle momenta of the process at hand. Aside from energy-momentum conservation, also the δ-functional angular resolutions must be taken into account. This means that the directions of particles must equal those of their corresponding reconstructed objects. In consequence, the phase space integrals to be performed in case of likelihood computations differ from those of cross section computations. Appropriate phase space generations are implemented. They allow for sufficiently fast computations. Details are given in App. D.

PDFs

The parton distribution functions are taken from the LHAPDF library [186]. In order to save computation time, all PDFs are cached before likelihood computations start. The factorization scale at which the PDFs are evaluated is set individually for each process.

Scattering amplitudes

The implementations of all amplitudes are taken from the MCFM program [187]. It provides a comprehensive collection of amplitude code which satisfies the needs of the present analysis. In particular, different descriptions of certain processes could be implemented, e.g. $t\bar{t}$ production with and without spin correlations or t-channel single

top-quark production in the 5-flavour and 4-flavour scheme. Also the amplitudes can be computed sufficiently fast. The relevant code was copied to the TELL library. Since MCFM is not specifically designed for this purpose, a few parts of the code were adjusted in order to create interfaces to the amplitudes. Cross checks are performed accordingly. For consistency, the strong coupling constant α_S is taken from LHAPDF and passed to the amplitude part of the code. In general, amplitudes must be distinguished by the charge of electrons or muons according to the measured event. Furthermore, there are processes which involve higher order corrections in terms of additional parton radiation. During the integration, the divergences involved in the corresponding amplitudes must be avoided. For this reason, cuts on minimum invariant masses of pairs of partons and minimum transverse momenta are applied accordingly. A list of the implemented scattering processes is provided in Sec. 3.2.3.

Transfer function resolutions

The resolutions of energies or transverse momenta are parametrised by double-Gaussian functions, which are taken from the KLFitter package [177] as provided within the ATLAS collaboration. They are derived from MC event simulations, where particles and matched reconstructed objects are compared with each other. A more detailed discussion of this subject follows in Sec. 3.2.4.

Transfer function efficiencies

The employed efficiencies provide the probability to detect a certain object given the momentum of its corresponding particle. There are reconstruction efficiencies as well as b-tagging efficiencies. They are determined using MC event simulations, similar to the case of the resolution functions. A more detailed discussion follows in Sec. 3.2.4.

Normalisation

In order to ensure the proper normalisation of the event likelihoods, the coefficients c_H must be determined for every scattering process H. Apart from that, they also depend on the topology of the measured event according to Eq. 3.27, namely on the numbers of jets, b-tags, electrons and muons. In the measurement at hand, the relevant topologies involve two b-tagged jets and one electron or muon. Accordingly, two normalisations

are determined for every process, depending on the lepton flavour. For this purpose, dedicated phase space generations are implemented. As no δ-functional resolutions enter these integrations, the involved phase space generations correspond to those which are typically used in cross section computations. In fact, if the efficiency term ε_H in Eq. 3.27 is neglected, the resulting integral equals the total cross section of the process H. This coincidence is used in order to cross check many parts of the TELL code. Indeed, the cross sections obtained from MCFM are reproduced successfully, which creates confidence in the implementation of the likelihood computation. [1]

Computer program

The elements introduced above are implemented and combined in terms of a program written in the C++ language. The only exception is given by the amplitude code of MCFM which is written in FORTRAN and integrated into the resulting library. The data analysis toolkit ROOT [188] is used in many parts of the code. Most notably, the ROOT file format known as "ntuples" is used for the input and output of analysis tasks. Hence the tool can be easily integrated into existing analysis frameworks.

The TELL code has a modular structure. Thus it is easy to maintain. In particular, extensions can be integrated quickly. The user interface is particularly simple. Still it provides a high degree of flexibility. In fact, the only work to be done by new users in order to integrate the code into their physics analysis is to program an event reader object which allows for the interpretation of the individual event format at hand. Due to the usage of the ROOT file format, this task can be accomplished straightforwardly. The user-friendliness of the tool is not only a matter of convenience. It constitutes a valuable aspect in view of the efficiency of experimental research, in particular in case of LHC experiments which need to deal with large amounts of experimental data and even more simulated events, given the numerous systematic variations under study.

An example of a program which configures a TELL computation is given below. The analysis is controlled with the help of a global management object. A collider, a PDF set and a set of transfer functions can be chosen as desired. The same is true for scattering processes, as shown by the choice of s-channel single top-quark production with two outgoing partons in the present case. The MC integration setup can be chosen

[1] There are a few cases in which this cross check is omitted, for instance s-channel single top-quark production with three final state partons. This is a higher order process and MCFM only allows one to integrate it as part of the full NLO cross section involving real and virtual corrections. The present TELL computation, on the other hand, does not involve virtual corrections.

for each process individually. Finally, the input of events and the output of likelihoods is defined, so that the computation can start.

```
Mgr *mgr = new Mgr;
mgr->SetCollider(Mgr::kPP, 8000.0);
mgr->SetPdfMgr("cteq66");

TFcnSet *tfcnATLAS = new TFcnSet(TFcnAtlasBase::kMC12);

ProcSgTop_sChannel_2jets *proc_sChannel_2j
            = new ProcSgTop_sChannel_2jets("sChannel2j",
                                           "SgTop s-channel, 2 jets");
proc_sChannel_2j->GetMCMgr()->SetEpsRel(5e-2);
proc_sChannel_2j->GetMCMgr()->SetMaxEval(5e5);
proc_sChannel_2j->GetMCMgr()->SetNStart(4000);
proc_sChannel_2j->GetMCMgr()->SetNIncrease(2000);
proc_sChannel_2j->SetTFcnSet(tfcnSetAtlasMC12);
mgr->AddProcess(proc_sChannel_2j);

mgr->SetEvtReader(new EvtReaderGeneric);
mgr->SetInputTreeName("t_evt");
mgr->AddInputFile("TellInput.root");

mgr->SetEvtWriter(new EvtWriterGeneric);
mgr->SetOutputFile("TellOutput.root");
mgr->SetOutputTree("t_llh", "TELL Likelihood Tree");

mgr->Run();
```

3.2.3. Scattering processes

There are eight different processes considered in the TELL computations of the present analysis. For each process H, the event likelihood $\mathcal{P}(X|H)$ is determined for every collision event X. Apart form the s-channel signal process, all major background contributions are taken into account. In the following, these processes are discussed. Table 3.6 summarises them and presents examples of corresponding Feynman diagrams. It is assumed that the top-quarks decay into a W boson and a b-quark.

***s*-channel single top-quark production, two final state partons:** This process models the signal contribution with its generic final state signature of a leptonic W boson decay and two b-quarks.

***s*-channel single top-quark production, three final state partons:** In order to improve the modelling of the signal contribution, this process is introduced additionally. It is a real radiation correction to the lowest order contribution mentioned above. Aside from two b-quarks, which are matched to the two b-jets of the measured event, the final state contains an additional light parton. It is not matched to any jet. Thus its momentum is integrated irrespective of any matching conditions. To avoid the soft and collinear divergences of the corresponding amplitude, the integration region is restricted to momentum configurations for which the invariant mass of this light parton and any of the other initial and final state partons is larger than 10 GeV. In addition, its transverse momentum must be larger than 10 GeV as well. The advantage of this process is the fact that it allows for finite transverse momenta of the system of the top-quark and the additional b-quark, thereby improving the modelling of the signal events.

***t*-channel single-top-quark production:** This process is modelled in the four-flavour scheme, where the initial state consists of a light quark and a gluon, which splits up into a $b\bar{b}$ pair. One of these b-quarks is matched to one of the measured b-jets. The same is true for the b-quark resulting from the top-quark decay. The other final state parton is integrated similarly to the case of the unresolved parton in the higher order s-channel single top-quark production described above. The finiteness of the width of the top-quark is neglected. This narrow-width approximation reduces the complexity of the MC integration, thereby speeding up the analysis.

***$t\bar{t}$* production, single-lepton:** Since most of the events under study originate from $t\bar{t}$ production, this process constitutes the most important background to be modelled. The two final state b-quarks are matched to the two b-jets of the measured event. Again, the narrow-width approximation is used in case of the top-quarks. Single-lepton $t\bar{t}$ production must be distinguished from the di-lepton case, as the transfer functions of these two processes are different. In the single-lepton process, the two light quarks resulting from one of the two W boson decays are integrated without any matching requirements.

$t\bar{t}$ production, di-lepton: In case of the di-lepton $t\bar{t}$ process it is assumed that one of the charged leptons resulting from the W boson decays is not reconstructed. In consequence, this lepton is integrated without any matching requirements.

W + 2 light partons: This process models the production of a W boson accompanied by two light partons, which are matched to the two measured jets. It is assumed that both of these jets are mis-tagged as b-jets.

$W + b\bar{b}$: The production of a W boson together with two b-quarks constitutes one of the most important background contributions. Its final state completely matches the topology of the measured events.

$W + c$ + light parton: The production of a W boson, a c-quark and a light parton and is considered as well. Both of these partons are assumed to be mis-tagged as b-jets.

3.2.4. Transfer functions

Within the TELL computations, all effects which take place subsequent to the hard scattering processes are modelled by means of approximate parameterisations. In particular, this includes the formation of jets and the detector response to the final state objects. Subsequent to the assignment of particles to measured objects, two effects are taken into account. These are finite reconstruction efficiencies and finite resolutions. They depend on the momenta of particles and reconstructed objects. These dependencies are determined with the help of simulated events where generated particles are compared with corresponding reconstructed objects.

Resolutions

Due to the high granularity of the ATLAS detector, the reconstruction of angles is approximately flawless. Finite energy resolutions remain and are parametrised by the functions $W^E_{\text{res}}(E^{\text{rec}}|E)$ introduced in Eq. 3.23. They are normalised with respect to E^{rec} according to Eq. 3.24. An exception is made in the case of muons, where the transverse momentum is used rather than the energy in order to obtain appropriate resolution functions. Here the coefficient $1/(E^{\text{rec}})^2$ in Eq. 3.23 must be replaced by $\sin\vartheta^{\text{rec}}/(E^{\text{rec}})^2$ since $\mathrm{d}^3p = E^2/\sin\vartheta\,\mathrm{d}p_{\mathrm{T}}\,\mathrm{d}\cos\vartheta\,\mathrm{d}\varphi$. In all cases the resolution is parametrised by a double Gaussian distribution. These functions are provided by the KLFitter package [177]

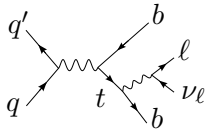
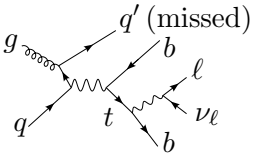
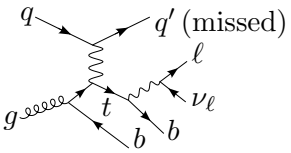
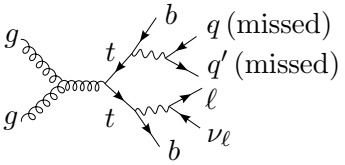
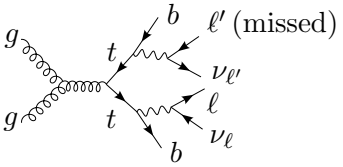
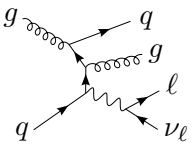
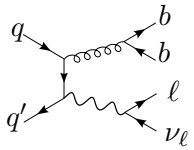
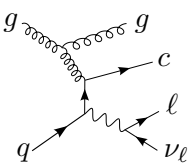
Process	Feynman diagram example
s-channel single top-quark production, two outgoing partons	
s-channel single-top-quark production, three outgoing partons	
t-channel single top-quark production, four flavour scheme	
$t\bar{t}$ production, single lepton	
$t\bar{t}$ production, di-lepton	
$W + 2$ light partons	
$W + b\bar{b}$	
$W + c +$ light parton	

Table 3.6.: Scattering processes considered in the TELL computations. In some cases, jets or leptons resulting from these processes do not meet the event selection conditions and are missed.

as distributed within the ATLAS collaboration,

$$W_{\mathrm{res}}^{E}(E_{\mathrm{rec}}|\,E) = \frac{1}{E}\frac{1}{\sqrt{2\pi}\,(p_2 + p_3 p_5)}\left(\exp\left(-\frac{(\Delta E - p_1)^2}{2p_2^2}\right) + p_3 \exp\left(-\frac{(\Delta E - p_4)^2}{2p_5^2}\right)\right), \tag{3.30}$$

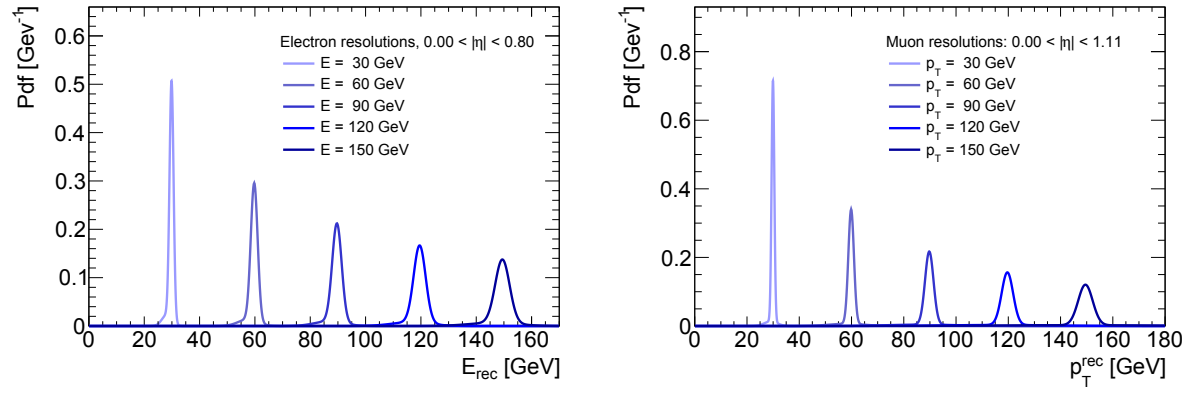
with ΔE being the relative energy difference

$$\Delta E = \frac{E - E^{\mathrm{rec}}}{E}\ . \tag{3.31}$$

The parametrisation of muon p_{T} resolutions is chosen accordingly. The parameters p_i depend on the particle's energy or transverse momentum. The particular parametrisations involve two constants per parameter. In most of the cases, a linear dependence $p_i = a_i + b_i E$ or $p_i = a_i + b_i p_{\mathrm{T}}$ is chosen. Exceptions are made in case of jet and electron energy resolutions, where the width p_2 of the first Gaussian term of Eq. 3.30 is parametrised according to $p_2 = a_2/\sqrt{E} + b_2$, in accordance with typical calorimeter energy resolutions.

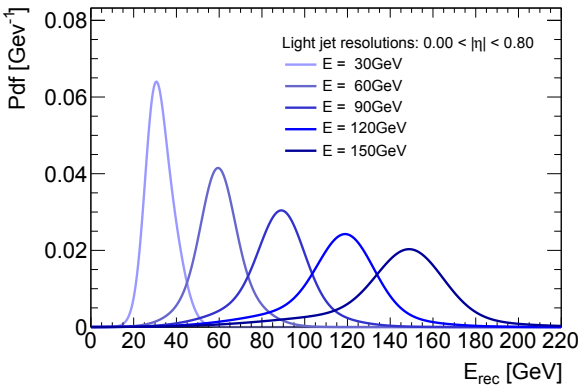
For each kind of measured object, namely light-flavour jets, b-jets, electrons and muons, the parameters a_i and b_i are determined in different regions of pseudorapidity based on a sample of simulated $t\bar{t}$ events. The sample is provided in App. A. Generated particles are considered if they match a corresponding reconstructed object, where the matching is defined by means of the angular distance, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} < 0.3$. In addition, the matching must be unique. For each matched pair, the relative energy difference ΔE is calculated. Muon p_{T} resolutions are treated accordingly. Finally, a fit of the resolution functions to the obtained distributions is performed in order to determine the parameters a_i and b_i. For this purpose a Markov-Chain Monte Carlo algorithm provided by the BAT program is used [189].

Figure 3.9 shows the resolution functions of physics objects for different particle energies or transverse momenta in the central detector region. Further resolution functions are provided in App. C.3. Clearly, the electron and muon resolution functions are much narrower than those of jets. In case of electrons and jets there is a bias towards lower reconstructed energies. In general, the widths of the resolution functions grow as the pseudorapidity increases. As an example, this is shown for the case of electrons in Fig. 3.9e.

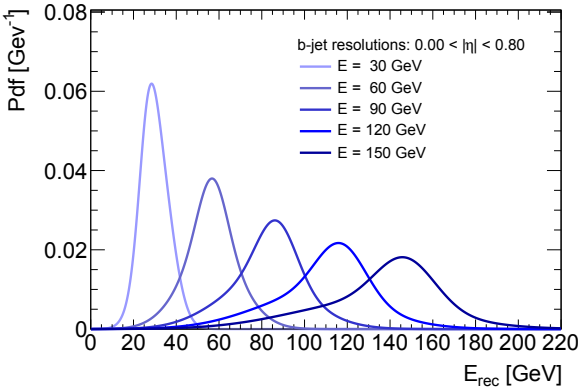


(a) Electrons

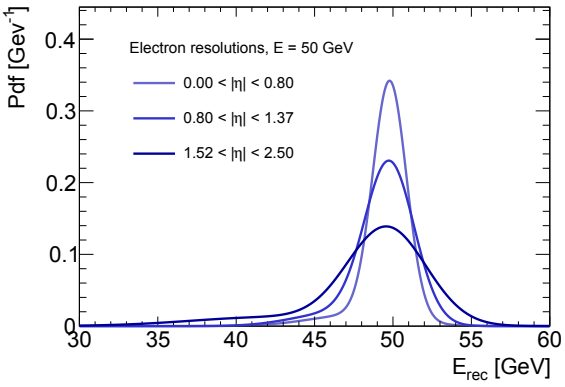
(b) Muons



(c) Light-flavour jets



(d) b-jets



(e) Electrons, different pseudorapidities

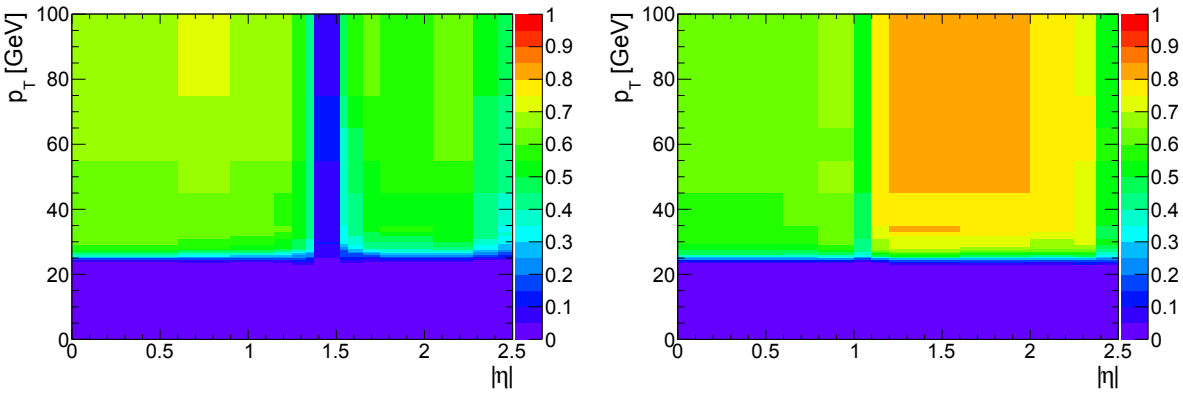
Figure 3.9.: Resolution functions of different physics objects. For central electrons (a), muons (b), light-flavour jets (c) and b-jets (d) the probability density of reconstructed energies is shown for different particle energies. In (e) it is shown for electrons as a function of the pseudorapidity given a constant particle energy.

Efficiencies

The efficiencies used in the transfer functions provide the probability of a final state configuration to pass the applied selection cuts. The overall efficiency is factorized into independent, single particle contributions which depend on the particle momentum at hand. There are reconstruction efficiencies of electrons, muons, light-flavour jets and b-jets, respectively, as well as efficiencies of the b-tagging of jets and of the applied $E_{\mathrm{T}}^{\mathrm{miss}}$ selection cut. The latter depends on the neutrino p_{T}. In case of the $t\bar{t}$ di-lepton process, where there are two neutrinos, the magnitude of the vectorial sum of the transverse neutrino momenta is used. The mis-tag efficiency is not determined but assumed to be approximately constant with respect to the parton momentum. The actual value of this constant mis-tag efficiency does not matter as it is cancelled once likelihoods are properly normalized according to Eq. 3.27.

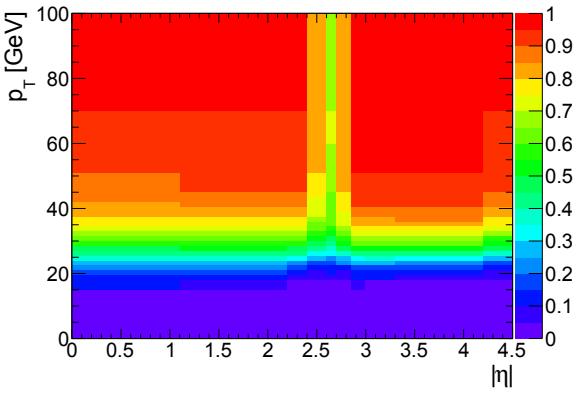
The efficiency determination is based on samples of simulated t-channel single top-quark events, which are provided in App. A. For every generated event, the top-quark decay products and the light quark produced in association with the top-quark are considered, which allows for the determination of all of the desired efficiencies. The event selection involves transverse momentum cuts of $p_{\mathrm{T}} > 25\,\mathrm{GeV}$. This threshold does not exactly match the event selection applied in the search for s-channel single top-quark production discussed above. However, like all approximations involved in the TELL computations, this difference affects the precision but not the validity of the final measurement. Matchings between particles and corresponding reconstructed objects are identified, where again a match is given if the angular distance between the two momenta is small, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} < 0.4$. The reconstruction efficiencies are given by the fraction of events where a particle is matched to a measured object which passes the corresponding object selection requirements. The b-tagging efficiency is given by the fraction of events where a reconstructed jet matched to the b-quark is b-tagged. Finally, the $E_{\mathrm{T}}^{\mathrm{miss}}$ efficiency is defined as the fraction of events which pass the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection cut.

In the context of the TELL computations, all efficiencies are used in the form of histograms as shown in Fig. 3.10 and 3.11. The transverse momentum cuts result in turn-on curves of the efficiencies with respect to p_{T}. In accordance with the momentum resolutions discussed above, the turn-on curves of electrons and muons are narrower than those of jets. Also characteristics of the ATLAS detector are present, namely reductions of several efficiencies caused by reduced amounts of instrumentation in certain regions of the detector. This affects the electron efficiency in the region $|\eta| \approx 1.5$, the muon

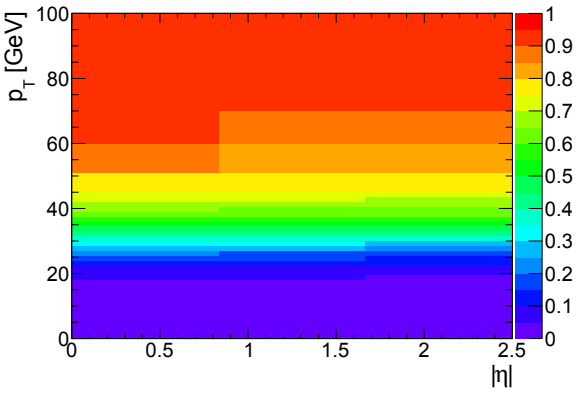


(a) Electron efficiency

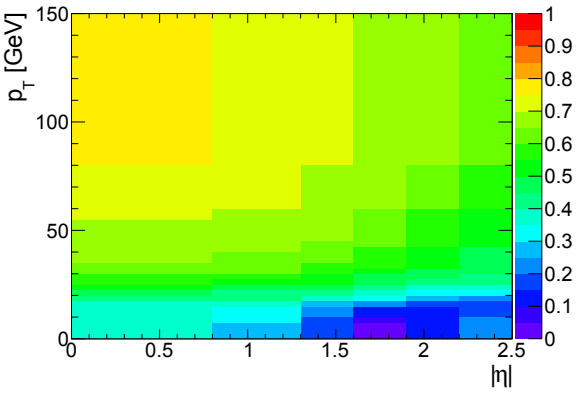
(b) Muon efficiency



(c) Light jet efficiency

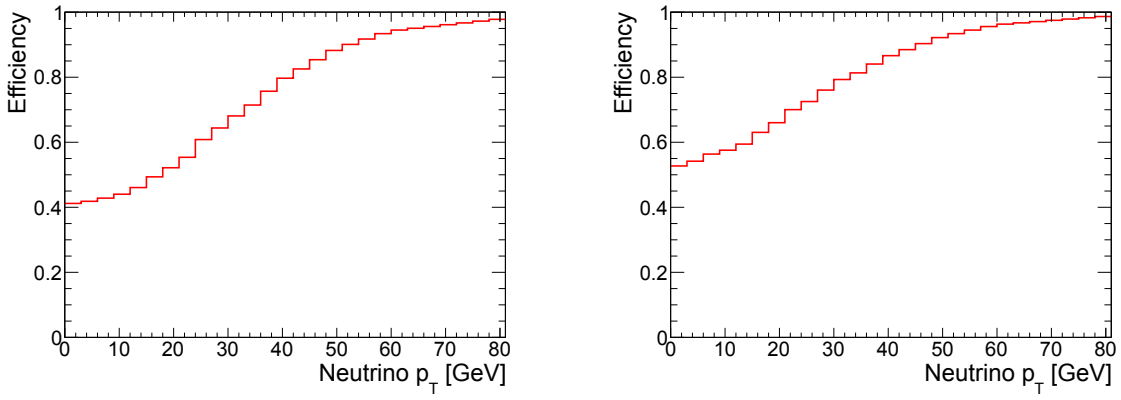


(d) b-jet reconstruction efficiency



(e) b-tag efficiency

Figure 3.10.: Transfer function efficiencies as a function of the pseudo-rapidity and the transverse momentum of primary particles. With respect to p_{T}, the reconstruction efficiencies (a) to (d) show turn-on curves around the cut of $p_{\mathrm{T}} = 25\,\mathrm{GeV}$ applied in the context of the efficiency study. The b-tagging efficiency shown in (e) reflects the chosen working point of about 70%. Distinct regions of reduced efficiency are caused by reduced amounts of instrumentation.



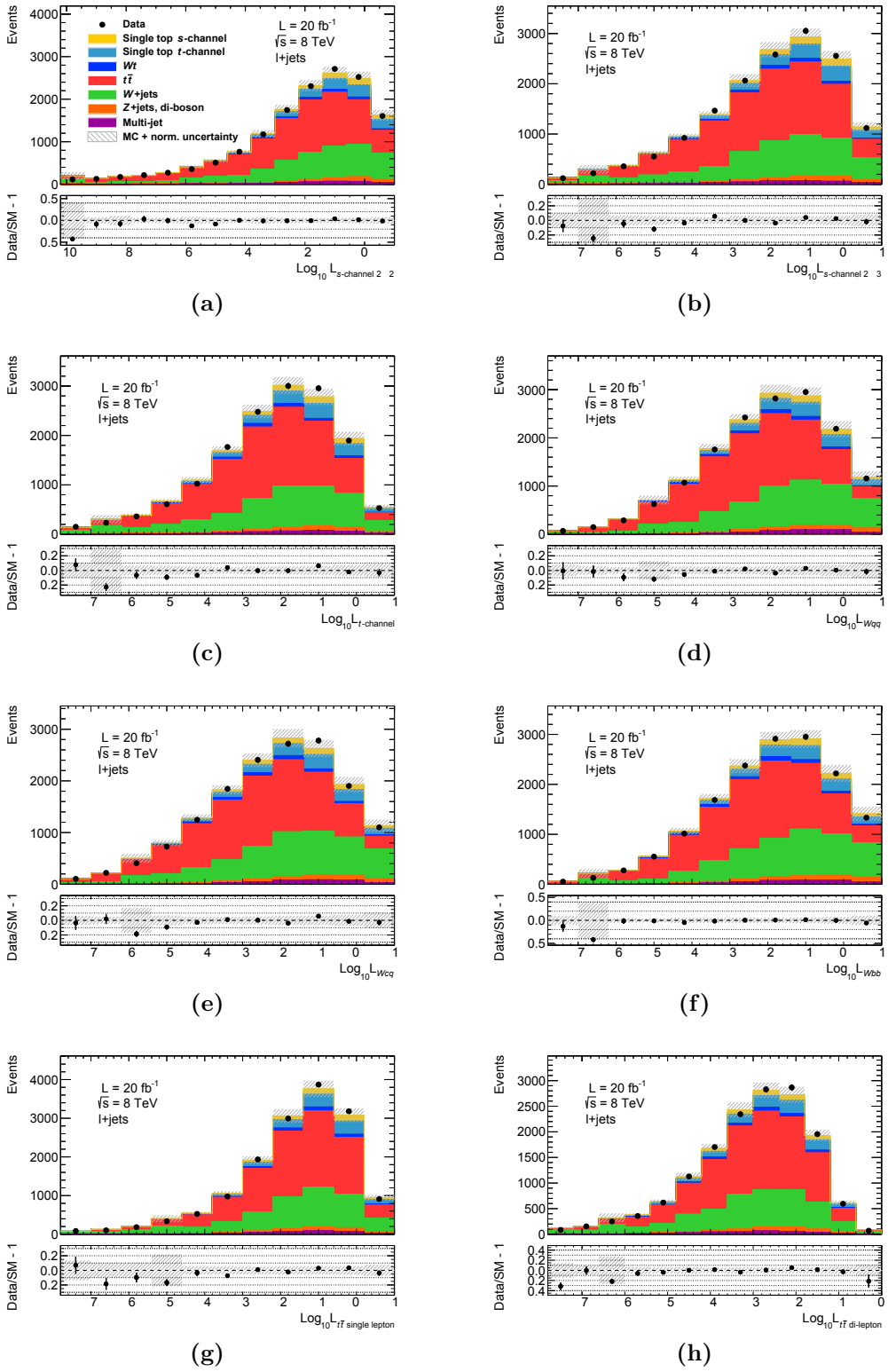
(a) Electron selection **(b)** Muon selection

Figure 3.11.: $E_{\mathrm{T}}^{\mathrm{miss}}$ selection cut efficiencies as a function of the neutrino transverse momentum in case the electron (a) and muon selection (b).

efficiency in the central region $|\eta| < 1.1$ and the jet reconstruction efficiency in the region $|\eta| \approx 2.5$.

3.2.5. Likelihood control distributions

Before a final discriminant is built and used in the search for s-channel single top-quark production, it must be checked that the distributions of TELLs which enter this discriminant are modelled correctly by the predictions in use. Figure 3.12 to 3.14 show the distributions of all TELLs in the signal region, in the W+jets control region and in the $t\bar{t}$ control region for the combined electron and muon selection. In case of the $t\bar{t}$ region, special care must be taken of the choice of jets used in the likelihood computations. While the events in question contain four jets, the likelihood input is restricted to two jets. Therefore only the two b-jets are used. Furthermore, they are required to be the jets with the largest transverse momenta while the other two jets must have $p_{\mathrm{T}} < 60\,\mathrm{GeV}$. These choices define a subset of the $t\bar{t}$ region which can be used in the likelihood computations and which is relatively similar to the signal region. In all regions a good overall agreement between the data and the prediction is achieved. These comparisons create confidence in the validity of the signal and background modelling in view of the final discriminant discussed in the next section.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 3.12.: TELL distributions in the signal region. Concerning the W+jets TELLs, q denotes light partons. The predictions are scaled according to the fit to the m_{T}^{W} distribution discussed in the previous section. The uncertainty bands correspond to the uncertainties due to the finite sample statistics and the theoretical normalisation uncertainties given after the fit.

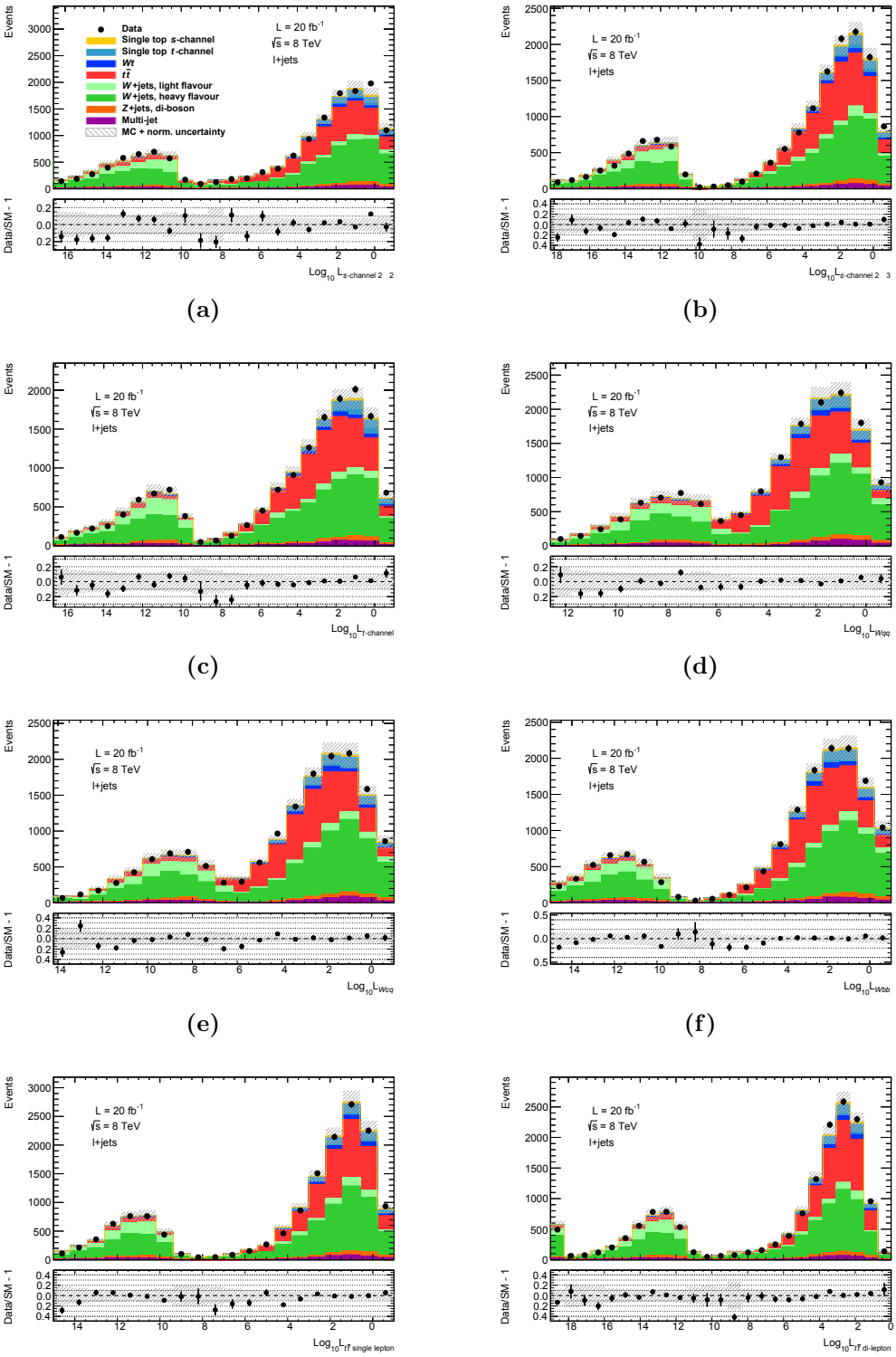


Figure 3.13.: TELL distributions in the W+jets control region. Concerning the W+jets TELLs, q denotes light partons. The predictions are scaled according to the fit to the m_{T}^{W} distribution discussed in the previous section. The uncertainty bands correspond to the uncertainties due to the finite sample statistics and the theoretical normalisation uncertainties given after the fit.

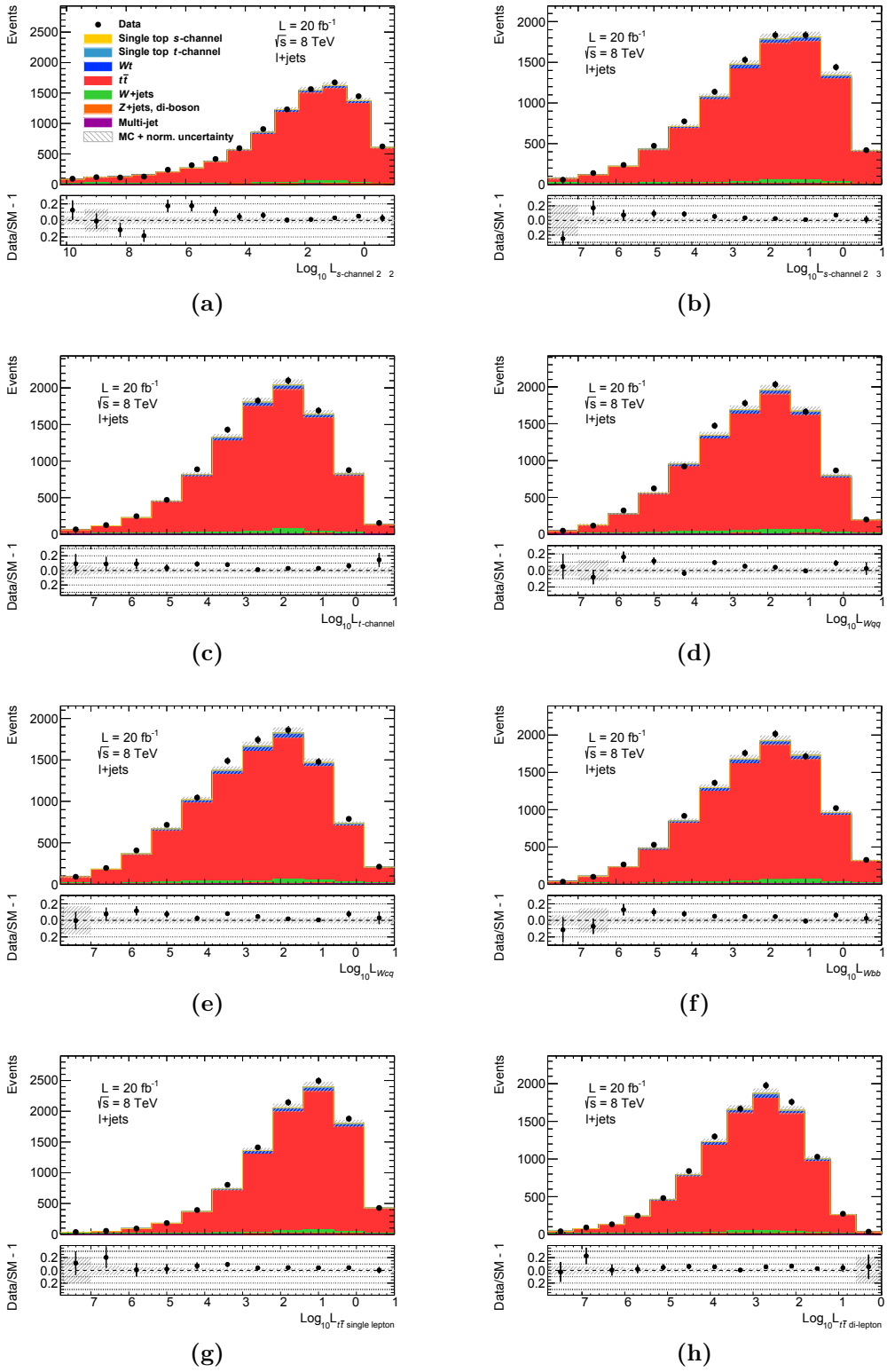


Figure 3.14.: TELL distributions in the $t\bar{t}$ control region. Concerning the W+jets TELLs, q denotes light partons. The predictions are scaled according to the fit to the m_{T}^{W} distribution discussed in the previous section. The uncertainty bands correspond to the uncertainties due to the finite sample statistics and the theoretical normalisation uncertainties given after the fit.

3.2.6. Event classification

Signal discriminant construction

The TELLs discussed above can be used in order to separate the signal from the background contributions. Single TELLs $\mathcal{P}(X|H_i)$, shown in Fig. 3.12 in case of the signal region, only provide relatively weak separations of processes. They need to be combined in order to optimise the separation power. Accordingly, the discussion of Sec. 3.2.1 led to the signal probability $P(S|X)$ as the optimal discriminant. The key formula Eq. 3.21 can now be specified with regard to the search for s-channel single top-quark production,

$$P(S|X) = \frac{\sum_{i\in\{S\}} P(S_i)\,\mathcal{P}(X|S_i)}{\sum_{i\in\{S\}} P(S_i)\,\mathcal{P}(X|S_i) + \sum_{j\in\{B\}} P(B_j)\,\mathcal{P}(X|B_j)} \ ,$$

$$\{S\} = \bigcup_{\ell\,\in\,\{e^+,e^-,\mu^+,\mu^-\}} \{S^\ell_{2\to2}, S^\ell_{2\to3}\} \ ,$$

$$\{B\} = \bigcup_{\ell\,\in\,\{e^+,e^-,\mu^+,\mu^-\}} \{B^\ell_{t\text{-channel}}, B^\ell_{t\bar{t}\,\text{single lepton}}, B^\ell_{t\bar{t}\,\text{di-lepton}}, B^\ell_{Wqq}, B^\ell_{Wbb}, B^\ell_{Wcq}\} \ . \quad (3.32)$$

The eight processes discussed in Sec. 3.2.3 are taken into account. Small contributions like Z+jets events and multijet events are neglected. The two signal processes, distinguished by the number of final state partons, are denoted by $S^\ell_{2\to2}$ and $S^\ell_{2\to3}$. All other processes enter the set of backgrounds $\{B\}$. Equation 3.32 underlines the fact that these processes are distinguished by the charge and flavour of the charged lepton resulting from a W boson decay. While the difference in lepton charge is a property of the hard scattering processes, the difference in lepton flavour is due to the transfer functions. Thus in general each of the eight likelihoods splits up into four different types. The only exceptions are the $t\bar{t}$ single and di-lepton likelihoods, which are charge–symmetric. Hence there are $6{\times}4{+}2{\times}2{=}28$ different processes in total.

Concerning the computation of TELLs, the dependencies on the lepton charge and flavour are taken into account automatically by the TELL program. Once all likelihoods are computed for a given event X, they must be combined with the corresponding a-priori probabilities $P(H_i)$. They are given by the fraction of events of the process H_i within the sample of all selected events as predicted by the simulations. The event

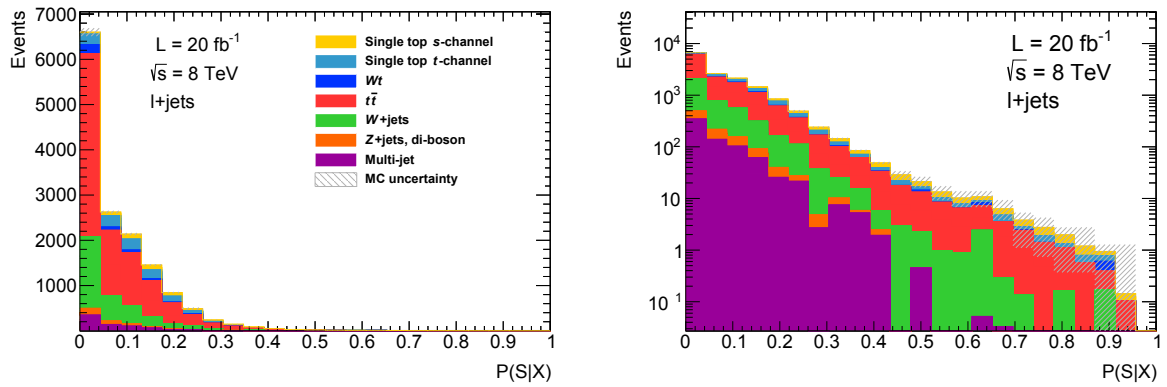


Figure 3.15.: Expected signal discriminant distribution defined by Eq. 3.32. A linear (a) and a logarithmic (b) event count axis scaling is shown. The s-channel signal is abundant at high discriminant values.

fractions were determined in a previous iteration of the analysis, published by the ATLAS collaboration [190]. Differences to the present analysis are given by the reference cross sections of single top-quark production. They are 6%, 3% and 1% larger than the cross sections used in the present analysis in case of the s-channel, t-channel and Wt production, respectively. Nevertheless the event fractions of the analysis published by ATLAS are used in the following. This choice makes a comparison of the two analyses easier. Furthermore, its impact on the final discriminant is small.

The partition of the s-channel processes with two and three final state partons is estimated with the help of the likelihood normalisation factors given in Eq.3.27,

$$\frac{P(s-\text{channel } 2 \to 2)}{P(s-\text{channel } 2 \to 3)} = \frac{\int \mathrm{d}\sigma^{2\to 2}_{s-\text{channel}}\, \varepsilon^{2\to 2}_{s-\text{channel}}}{\int \mathrm{d}\sigma^{2\to 3}_{s-\text{channel}}\, \varepsilon^{2\to 3}_{s-\text{channel}}} \ , \tag{3.33}$$

which leads to a ratio of 0.65/0.35 in favour of the 2→2 process. The W boson + heavy flavour jets background is associated with the Wbb and Wcq likelihoods, where a relative fraction of 80% is chosen in favour of the more prominent Wbb process.

With these entities, the approximate signal probability $P(S|X)$ can be constructed. It is the input to the statistical evaluation of the measured data discussed below in Sec. 3.4. Up to this point, only the expected distribution is discussed in order to focus on the interpretation of the discriminant first. Also the validity of the modelling of this distribution is investigated first using the two control regions. The measured distribution

in the signal region is presented when the s-channel single top-quark measurement is discussed in Sec. 3.4.

Figure 3.15 shows the expected distribution of the discriminant given the MC simulations and lepton fake estimate. The presented versions of the distribution differ by the scaling of the event count axis being linear or logarithmic. As expected, s-channel single top-quark events accumulate at high discriminant values. This separation allows for a search for this process given the measured data. However, the histograms constructed from these distributions using an equidistant binning are not necessarily suited for this purpose. Therefore, an alternative binning is determined. The rebinning is performed in the context of the s-channel measurement, maximising the expected significance of the signal process. This results in another discriminant, denoted by "Transformed $P(S|X)$" in the following, which is a monotonic function of the original discriminant $P(S|X)$. Its distribution is shown in Fig. 3.16. Details of the rebinning procedure are provided in App. C.4. Furthermore, the new distribution is restricted to a subset of the available events by means of a cut which excludes the lowest discriminant values. This cut is suggested by studies of the modelling of the discriminant distribution, which are presented in App. C.4 as well.

Discriminant control distributions

Despite the various investigations of the goodness of modelling of the data by means of kinematic distributions as well as TELL distributions presented above, also the modelling of the final discriminant $P(S|X)$ is investigated in order to create more confidence in the employed simulations and lepton fake estimates. Figure 3.17 presents the distribution of this discriminant in the W+jets and $t\bar{t}$ validation regions, respectively. Again, in case of the $t\bar{t}$ region a subset of events is chosen in order to be able to compute the TELLs as explained in Sec. 3.2.5. Overall a good agreement between the data and the prediction is achieved.

3.2.7. Further studies

The discussion of the TELL method presented above is sufficient in order to proceed to its application in the context of a measurement. In addition to that, further aspects of the method are discussed throughout the rest of this section.

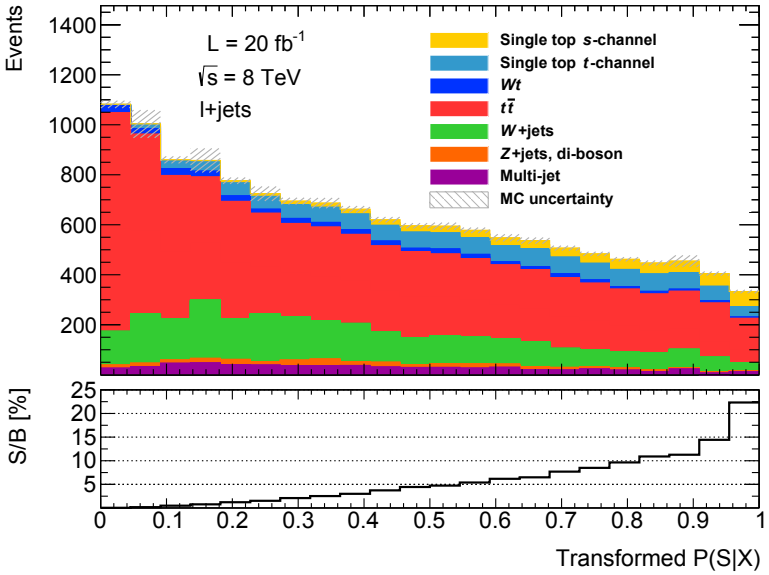
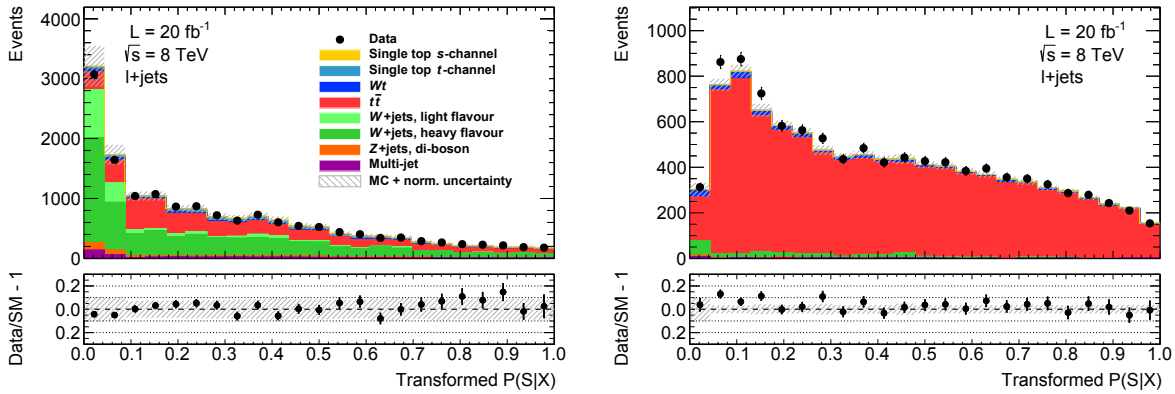


Figure 3.16.: Final expected signal probability distribution. The original distribution $P(S|X)$ is mapped monotonously onto the new distribution, which is better suited for a search for the s-channel signal process in the context of binned maximum likelihood fits. Particularly low discriminant values are neglected when this final discriminant is built. The lower panel shows the ratio of signal and background event counts in each bin.



(a) W+jets region **(b)** $t\bar{t}$ region

Figure 3.17.: TELL discriminant control distributions. All templates are scaled according to the control distribution fit results discussed in Sec. 3.1.4.

Computational performance

The time needed for the computation of the TELLs is crucial to the applicability of this method. In contrast to machine learning algorithms, which can be applied quickly, the

Process	Precision [%]	Time per event [s]	$1 - \frac{\text{SEP}_{\text{no proc i}}}{\text{SEP}_{\text{all}}}$ [%]
s-channel $2 \to 2$	13	1.9	8.1
s-channel $2 \to 3$	27	7.6	2.3
t-channel	61	8.7	0.41
$t\bar{t}$, single lepton	16	9.1	8.6
$t\bar{t}$, di-lepton	18	9.2	0.37
Wqq	9.2	2.7	0.55
Wcq	9.3	5.0	0.29
Wbb	10	0.85	1.4

Table 3.7.: Performance of TELL computations. The precision of the MC integration is determined by means of comparisons between the nominal and higher precision integrations based on the same set of events. For this purpose a set of measured events is used in order to take into account all relevant scattering processes. The same is true in case of the evaluation of the computation time. The rightmost column presents the impact of each process on the signal separation power defined by Eq. 3.34.

integration of TELLs is rather time-consuming. This is particularly important if there are many systematic uncertainties which require dedicated TELL computations due to variations of momentum components.

Table 3.7 summarises the computational performance. First of all it presents the statistical precision of the MC integration reached for each process. Mostly it is on the order of 10%, which is a compromise between time consumption and precision. Only in case of the t-channel single top-quark process, the precision is rather low. The time needed for a single TELL computation varies between about one second and ten seconds, depending on the kinematic configuration of the process at hand. Thus the complete set of necessary TELL computations in the given analysis can be completed within a relatively short time. Using a combination of the Worldwide LHC Computing Grid and a local computer cluster as explained in Sec. 3.1, about two weeks are needed in order to perform the computations. This is a competitive performance. It is achieved by means of thorough studies of the TELL code and by appropriate phase space parametrisations as explained in App. D.

Relevance of transfer functions

Studies of transfer functions are one of the main efforts necessary to implement the TELL method. Thus it is of interest to determine their impact on the performance of the method. For this purpose, different transfer function setups are compared with each other with regard to the resulting separation of the signal. In order to quantify the separation power, cuts on minimal values of the TELL discriminant are applied. This results in the background rejection as a function of the signal selection efficiency. For each cut value $P_{\min}(S|X)$, all events with $P(S|X) > P_{\min}(S|X)$ are considered. The background rejection r_{bkg} refers to the sum off all background events and is related to the background selection efficiency ε_{Bkg} via $r_{\text{bkg}} = 1 - \varepsilon_{\text{Bkg}}$. Summarising the relation between r_{Bkg} and the signal selection efficiency ε_{Sig}, the resulting curve provides a measure of the signal separation, known as the Receiver-Operating-Characteristic (ROC). Hence large integrals of this curve indicate large background rejections for a given signal efficiency. Since this integral is bounded below by one half anyway, the following measure is used to define the signal separation, yielding values limited by zero and one,

$$\text{SEP} = 2 \int_0^1 \mathrm{d}\varepsilon_{\text{Sig}}(1 - \varepsilon_{\text{Bkg}}) \; - \; 1 \; . \tag{3.34}$$

Figure 3.18 shows the ROC curves resulting from three different sets of transfer functions. Apart from the chosen set of transfer functions, the discriminant is constructed in the same way as discussed above. The ROC curve resulting from the choice of the nominal set of transfer functions as applied in case of Fig. 3.16 serves as the reference. If selection efficiencies are neglected, the signal separation is reduced by 1.4%. Thus the analysis benefits from their inclusion, but only by a small amount. Certainly an optimisation of the efficiency study where its p_{T} selection thresholds exactly match those of the final event selection could increase the signal separation further.

The impact of transfer function resolutions is estimated by a comparison with a simple set of resolution functions. The double Gaussian functions parametrising the resolutions of the ATLAS detector are replaced by single Gaussian distributions. A width of 4% is used in case of electron and muon resolutions, whereas the width of jet resolutions is chosen to be 25%. These values approximate the resolutions of the ATLAS detector. However, dependencies of these widths on particle energies are neglected. As a result, the signal separation is reduced by 5.3%. Hence the usage of detailed parametrisations

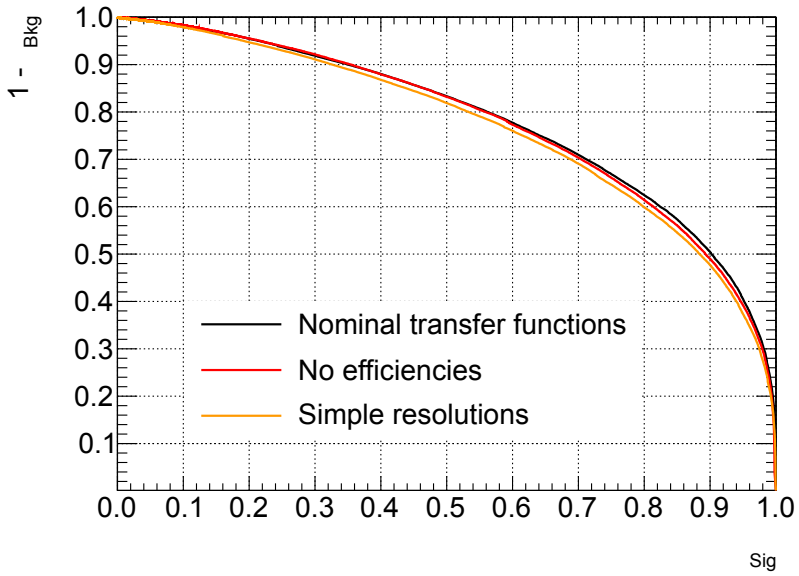


Figure 3.18.: Impact of transfer functions on the separation power of the TELL discriminant. ROC curves resulting from cuts on minimal values of the discriminant allow one to compare the impact of different choices of transfer functions on the signal separation. The TELL method benefits from the inclusion of selection efficiencies and detailed parametrisations of resolutions.

of resolution functions which are in accordance with the detector response improves the separation power noticeably.

Relevance of single processes

Table 3.7 also shows the impact of each TELL on the signal separation. For each process the rightmost column gives the relative reduction of the signal separation which is obtained if the process in question is not taken into account in the construction of the final discriminant. A large reduction indicates a large impact of the particular process on the signal separation.

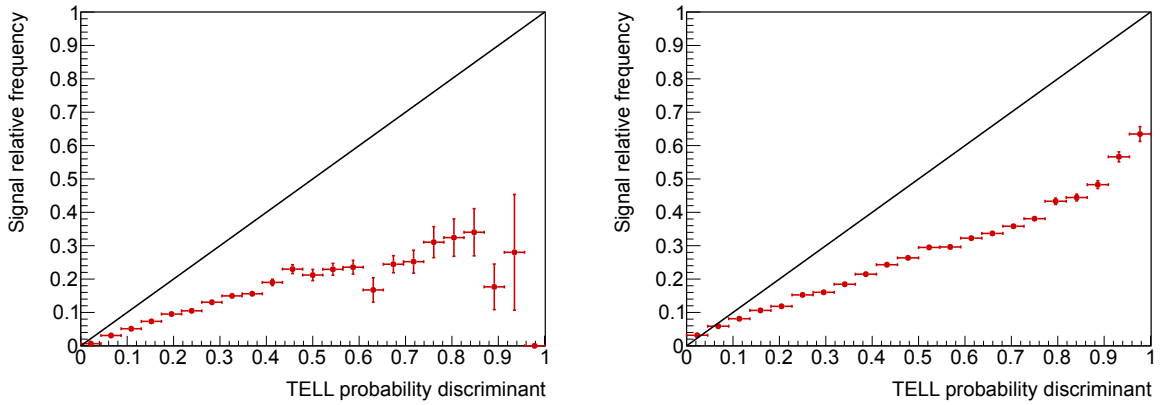
The most important processes are the s-channel signal with two outgoing partons and the $t\bar{t}$ single lepton background. The s-channel signal process with three outgoing partons and the Wbb process are important too. The $t\bar{t}$ di-lepton likelihood and the t-channel single top-quark process have a very small impact on the signal separation even though they are major backgrounds. In the latter case, the small impact is assumed to be related to the low precision of the MC integration of this TELL. In case of the

$t\bar{t}$ di-lepton process the small impact suggests that the corresponding TELL can be improved considerably.

Goodness of likelihood approximations

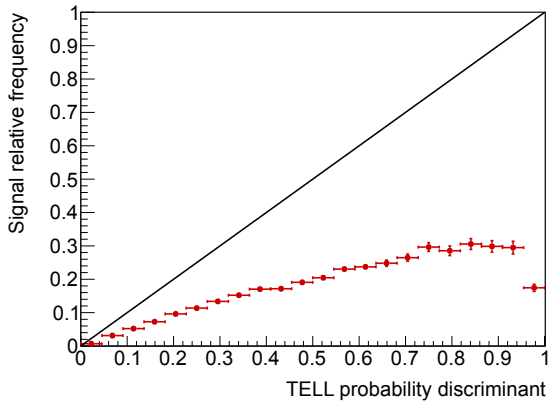
The validity of the approximations which enter the TELL computations is studied in the following. The approximate signal probability $P(S|X)$ based on the computed TELLs is compared with the fraction of signal events in each bin of the discriminant distribution as predicted by the simulations and lepton fake estimates. This is shown in Fig. 3.19a. In case of a perfect modelling of event probability densities by the TELLs, the signal event fraction equals the discriminant $P(S|X)$. This scenario is indicated by the black diagonal line, whereas the signal event fraction predicted by the simulations and lepton fake estimates is given by the red points. The latter values increase linearly with the discriminant value as expected. On the other hand, these values turn out to be smaller than those given in the optimal case. This observation reveals that the computed TELLs do not match the event probability densities perfectly.

A closer look at this deviation can be taken with the help of Fig. 3.19b to 3.19d. In these cases only single background processes are taken into account in the construction of the discriminant $P(S|X)$. Accordingly, the provided signal event fraction only refers to the background events in question. In case of the single top-quark t-channel background, Fig. 3.19b, a deviation with respect to the optimal case is observed. This shortcoming can be related to the relatively low precision of the MC integration of t-channel TELLs discussed above. In case of the $t\bar{t}$ background, Fig. 3.19c, this deviation is even larger. In particular, the signal event fraction decreases at high discriminant values, contrary to the expectation. Together with the observation of a small impact of $t\bar{t}$ di-lepton TELLs on the signal separation discussed above, this finding suggests that the implementation of this TELL can be improved by a large amount. Finally, Fig. 3.19d presents the same study with respect to the W+jets background. In this case, the observed signal event fraction matches the expectation well. This observation suggests a good modelling of event probabilities by the TELLs with respect to the signal as well as the W+jets hypotheses. Overall, this study shows that the employed implementation of the TELL method can still be improved considerably, despite the fact that it separates the signal process from its backgrounds successfully.

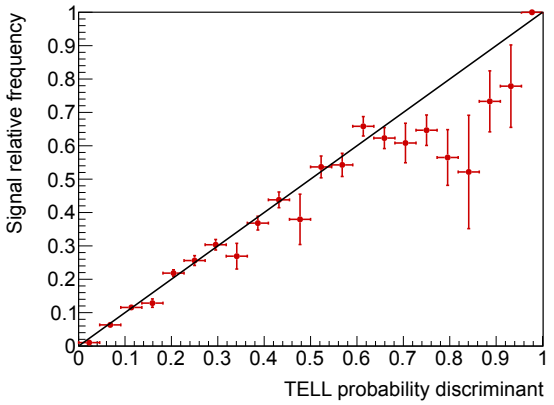


(a) All backgrounds

(b) single top t-channel



(c) top-quark pair production



(d) W+jets

Figure 3.19.: Approximation of signal probabilities by TELL discriminants. Each figure refers to another set of background processes taken into account in the construction of the TELL discriminant $P(S|X)$. The red points show the signal event fraction given the background contributions in question as predicted by the employed simulations and lepton fake estimates. The black diagonal lines refer to the scenario of an optimal modelling of event probability densities by TELLs.

3.3. Systematic uncertainties

The modelling of the TELL discriminant distribution is subject to various uncertainties. There are statistical uncertainties due to quantum mechanical fluctuations and systematic uncertainties. All of them must be taken into account in order to be able to estimate the number of signal events present in the measured data set. This estimation is performed with the help of a statistical model, which incorporates the uncertainties. In particular, the systematic uncertainties are parametrised in terms of probability densities. They become part of a likelihood function which allows one to determine the signal strength. The mean values of these densities are variables, called nuisance parameters. In the following, the sources of systematic uncertainties are discussed. Concerning instrumental uncertainties, the discussion refers to Sec. 2.5 where the reconstruction of the relevant objects is explained together with corresponding uncertainties. The details of the statistical evaluation, on the other hand, are discussed below in Sec. 3.4.

3.3.1. Sources of uncertainties

On the one hand, the limited number of simulated events as well as measured events regarding the prompt lepton fake estimate causes statistical uncertainties of the predictions. Further systematic uncertainties arise from finite instrumental precision and limited abilities to model scattering processes.

Instrumental uncertainties

Luminosity

The data set in use corresponds to an integrated luminosity of $20.3\,\mathrm{fb}^{-1}$, which is determined with the help of beam profile scans as outlined in Sec. 2.3. The uncertainty of this measurement amounts to 2.8%. All samples of simulated events are normalised on the basis of this luminosity measurement using Eq. 3.2 and its uncertainty is taken into account accordingly.

Jet energy scale

As discussed in Sec. 2.5.4, the energy calibration of measured jets consists of two steps. Firstly, the detector response to jets beyond the central region $|\eta| < 0.8$ is related to the response given in case of central jets by means of the dijet η-intercalibration technique. Apart from statistical uncertainties, the η-intercalibration is subject to uncertainties of the modelling of jets at high rapidities, where predictions of the employed event generators, namely Pythia and Herwig, differ.

Secondly, the absolute jet energy scale is set on the basis of the Z+jet, γ+jet and multijet balance techniques, where the transverse momentum of the jet is related to a well-calibrated reference object. The difference between the resulting ratios ${p_\mathrm{T}}^{\mathrm{jet}}/{p_\mathrm{T}}^{\mathrm{ref}}$ in data and simulation leads to the energy correction factor $\mathcal{R}({p_\mathrm{T}}^{\mathrm{jet}}, \eta^{\mathrm{jet}})$ defined by Eq. 2.10. A combination of the $\mathcal{R}$ ratios resulting from the three balance techniques gives the final calibration factor. There are several uncertainties which are associated with the three single methods. They are related to the reconstruction of physics objects, like e.g. the photon energy scale, to the modelling of scattering processes, like e.g. event simulations based on different MC event generators, and to finite amounts of statistics. These uncertainties are propagated to the combined result. Some of them turn out to have only a small impact on the jet energy scale. In order to reduce the complexity of statistical models, the number of uncertainties is reduced by means of combinations of nuisance parameters associated with uncertainties of small significance.

Furthermore, the jet energy scale is affected by uncertainties related to the subtraction of contributions caused by pile-up events. The pile-up correction depends on the p_T area density of particles produced in soft pp interactions, on the mean number of primary vertices, on the number of primary vertices in the collision event at hand and on the jet p_T. Uncertainties related to all of these dependencies are taken into account.

Another source of uncertainty on the jet energy scale is related to the types of partons inducing a jet. In general the relevant parton flavour fractions differ between event samples used for physics analyses and those used for the jet calibration. As the jet p_T depends on the type of the inducing parton, this difference causes an uncertainty, which is referred to as the "flavour composition" uncertainty. Furthermore, the modelling of jets induced by gluons is performed in a significantly different way by different event generators. The corresponding uncertainty is referred to as the "flavour response" uncertainty [191]. An additional uncertainty of the calorimeter response to b-jets is taken into account as well [155]. Furthermore, the uncertainty of the energy leakage beyond the

calorimeter, known as "punch through", is taken into account. In Tab. 3.8 jet energy scale uncertainties are denoted by "Jes", followed by the name of the particular uncertainty component.

Further systematic uncertainties of jets

The energy resolutions of jets measured in data on the one hand, and of simulated jets on the other hand, agree with each other to a large extent, as outlined in Sec. 2.5.4. Residual differences are taken into account as a systematic variation, where an additional smearing is applied to simulated jets. The resulting single variation of the final discriminant is symmetrised. Concerning the jet vertex fraction, the corresponding cut is varied upwards and downwards, which results in two variations. They cover the discrepancy between the impact of this cut on measured and simulated jets. The jet reconstruction efficiency is close to unity in case of the jets considered in the present analysis. The determination of this efficiency makes use of jets based on tracks of charged particles, which are matched to jets based on calorimeter cells [192]. Residual discrepancies between the efficiency obtained in data and simulations are taken into account by randomly discarding jets in simulated events at an appropriate rate. Again, the resulting variation of the final discriminant is symmetrised.

Jet flavour tagging

The calibration of the b-tagging algorithm described in Sec. 2.5.4 is subject to various uncertainties. They are considered for the three cases of actual b-jets, charm jets and light-flavour jets, which results in three uncorrelated uncertainties, called the b-tagging, c-tagging and mis-tagging uncertainty. The most important tagging uncertainty with regard to the present analysis is the b-tagging uncertainty. The impacts of the single sources of its uncertainties are summed up in quadrature, which results in upward and downward variations of the b-tagging scale factors depending on the transverse momentum of the jets. Apart from statistical uncertainties, both instrumental and modelling uncertainties are taken into account. These systematic uncertainties overlap with the uncertainties taken into account explicitly in the present analysis, which introduces a double counting. However, the dominating systematic uncertainties of the b-tagging calibration are related to the modelling of top-quark pair production and to the jet energy scale calibration,

which turn out to be minor uncertainties in the present analysis. For these reasons, the double counting of uncertainties is negligible [161].

Electrons and muons

Further uncertainties arise from the calibrations of electron and muon momenta. In both cases, uncertainties of the trigger efficiency are taken into account. The same is true for the bare electron and muon reconstruction efficiencies. Further uncertainties given in case of electrons are related to the electron energy scale and to the electron energy resolution. In the latter case, an additional energy smearing is applied to simulated electrons in order to match the resolution obtained in data. This smearing is varied. Also the uncertainty of the electron identification efficiency is estimated. Concerning muons, there are uncertainties related to the scale and resolution of their momentum. In the latter case, dedicated uncertainties of inner detector and muon spectrometer momentum resolutions are taken into account. Only single variations are performed in these cases and their effect on the final discriminant is symmetrised, which leads to up and downward variations. The uncertainty of the muon isolation requirement is considered as well.

Missing transverse momentum

On the one hand, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is affected by uncertainties of jets and leptons. They are taken into account as described in the previous paragraphs and are propagated to the corresponding terms entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. On the other hand, uncertainties of the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term are taken into account additionally. The scale and the resolution of this contribution are varied up and downward.

Modelling uncertainties

Background normalisation

The simulated scattering processes are normalised according to higher order cross section calculations using Eq. 3.2, taking into account the corresponding uncertainties. In case of the t-channel single top-quark production and Wt associated production backgrounds, these uncertainties amount to 4% and 6%, respectively. These values result from scale

uncertainties presented in App. B and from PDF uncertainties given in Ref. [72]. In the context of control distribution fits, where the s-channel signal is constrained by its SM prediction, an uncertainty of 5% is taken into account in accordance with Sec. 1.3.4. The uncertainty of the top-quark pair production normalisation amounts to 6%, as discussed in Sec. 1.3.2.

In case of the W+jets normalisation, multiple uncertainties are summed up in quadrature. The uncertainty of the total cross section of inclusive W boson production amounts to 4% [166]. With regard to the two selected jets, an uncertainty of 24% is taken into account twice, in accordance with the Berends scaling rationale [193, 194]. Concerning the fraction of heavy flavour jets contained in the W+jets event sample, an additional uncertainty of 50% is taken into account. Overall, these contributions add up in quadrature to an uncertainty of 60%. This low precision estimate leaves the W+jets normalisation weakly constrained, so that it is determined more precisely by the fit of the statistical model to the measured data. The Z+jets normalisation uncertainty is determined in the same way, which also results in an overall uncertainty of 60%. The di-boson background is merged with the Z+jets contribution. Both of these processes are backgrounds of minor importance. The normalisation of the lepton fake background is varied by 50%, which covers the uncertainty of this data-driven estimate as discussed in Sec. 3.1.2 [168].

Event simulation

Concerning the simulation of the relevant scattering processes, different event generators are available. In addition to the baseline event generator choice, alternative generators are employed in order to estimate the robustness of the final results with regard to modelling uncertainties. The particular generators are mentioned above in Sec. 3.1. Each generator comparison leads to a single variation of the discriminant distribution. These variations are symmetrised, so that there are upward and downward variations.

Furthermore, renormalization and factorization scales are varied in case of the s-channel signal and top-quark pair production background, which is the dominant process in the present analysis. In the latter case, the hardness of QCD radiation is varied additionally by means of a dedicated parameter, known as the h_{damp} parameter. For these uncertainties, the maximum and minimum of the discriminant in each histogram bin is chosen in order to construct upward and downward variations. In case of the

top-quark pair production background, this uncertainty is denoted by "$t\bar{t}$ modelling" in the following.

In general, the way in which these modelling uncertainties are evaluated is rather simple. In the context of the statistical model, nuisance parameters are used to interpolate between discriminant distributions, even though all values apart from those predicted by a certain model cannot be interpreted properly. A rigorous approach to these uncertainties would require detailed studies of a lot of aspects associated with the simulation of collision events. This is beyond the scope of the present analysis. Instead, the simple approach described above is chosen in order to figure out if these uncertainties have only a small impact on the final result, so that there is no need for more detailed modelling studies. This condition is met in case of most of the generator comparisons of the present analysis. The only exception is given by the case of the modelling of $W+$jets events, where the Sherpa generator is compared with Alpgen interfaced with Pythia 6. The choice between these generators can have a significant impact on the final results. A discussion of this issue is given below in the context of discriminant shape variations due to systematic uncertainties. It is based on a detailed comparison of the two $W+$jets predictions, which is presented in App. C.5.

Parton distribution functions

Uncertainties of the proton PDFs are propagated to the discriminant distributions by means of a reweighting of events, where the event weight is adjusted by ratios of varied over nominal PDFs, evaluated at the parton momentum fractions x_1 and x_2 of the collision event at hand. The evaluation of the PDF uncertainty is based on the PDF4LHC prescription discussed in Sec. 1.3.4 [59]. This leads to envelopes of discriminant distributions, denoted by $P_{\text{PDF}}^{\text{high}}$ and $P_{\text{PDF}}^{\text{low}}$ in the following. The procedure is performed for all simulation samples. Some of the baseline generators are replaced by other generators in order to be able to determine the parton momentum fractions x_1 and x_2 for each event. In these cases, the alternative generator only serves the determination of the PDF uncertainty, which is then used as a variation with respect to the nominal discriminant distribution using the baseline generator. Finally, the PDF variations incorporated in the statistical model are constructed as follows, where P_{nominal} denotes the nominal

distribution of the discriminant,

$$P_{\mathrm{PDF}}^{\mathrm{up/down}} = \left(1 \pm \frac{P_{\mathrm{PDF}}^{\mathrm{high}} - P_{\mathrm{PDF}}^{\mathrm{low}}}{P_{\mathrm{PDF}}^{\mathrm{high}} + P_{\mathrm{PDF}}^{\mathrm{low}}}\right) \cdot P_{\mathrm{nominal}} \quad . \tag{3.35}$$

3.3.2. Impact on rates

Table 3.8 presents the upward and downward variations of the total rate of each process caused by each of the uncertainties in case of the signal region. The sources of uncertainties with the largest impact on rates of several processes are the b-tagging and the choice of the MC event generator. The lepton fake background is not listed because this data driven estimate is only subject to its 50% normalisation uncertainty.

3.3.3. Impact on discriminant shapes

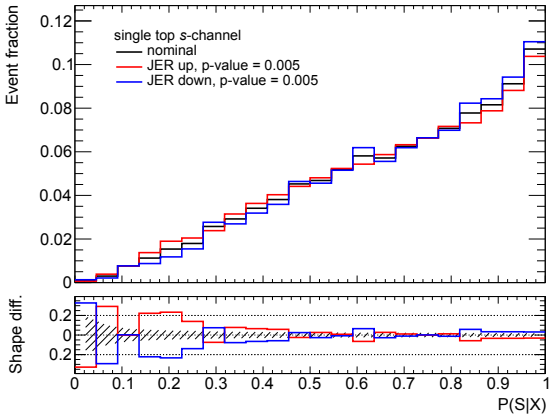
Most of the sources of systematic uncertainties affect the shape of the discriminant distribution. The only exceptions are global rate uncertainties, namely the luminosity uncertainty and the background normalisation uncertainties caused by the finite precision of reference cross section calculations. Still, many of the shape variations caused by the systematic uncertainties are smaller than the fluctuations related to finite simulation statistics. In these cases, only the variation of the overall rate is taken into account. This choice is made to prevent a double counting of uncertainties.

In order to determine the sources of uncertainties which have a significant impact on the shape of discriminant distributions used for the signal extraction, χ^2-tests are performed. For each of the simulated processes and for each of the systematic uncertainties the systematically varied distributions are compared with the corresponding nominal one. The varied distributions are renormalised according to the normalisation of the nominal distribution as these tests only serve the identification of significant shape variations, while overall rate uncertainties are considered anyway.

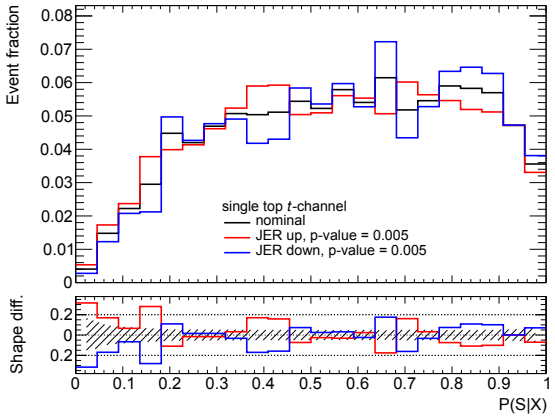
The nominal and systematically varied distributions are derived from disjoint sets of events. The splitting makes use of the event numbers, which are assigned to each event individually during the simulation process. The nominal distributions are derived from events with an even number, whereas odd event numbers are used to derive the

Rate uncertainty [%]	s-channel		t-channel		Wt		$t\bar{t}$		W+jets		Z+jets & VV	
Source	up	down	up	down	up	down	up	down	up	down	up	down
η-intercalib. stat.	-0.1	0.1	-0.8	0.3	-1.4	1.3	-0.9	1.0	-0.6	-0.3	-0.4	-0.5
η-intercalib. model	-0.3	0.2	-2.0	1.5	-1.0	1.4	-1.3	1.5	-0.7	-0.3	-0.8	0.3
Jes eff. det. 1	0.1	-0.1	-0.0	-0.2	-1.0	0.8	-0.2	0.3	-0.2	-0.1	-0.1	0.0
Jes eff. det. 2	-0.0	0.0	-0.1	-0.1	-0.5	0.9	-0.2	0.3	-0.4	-0.1	-0.4	-0.1
Jes eff. det. 3	0.0	0.0	-0.1	-0.1	-0.4	-0.2	0.0	0.1	-0.4	-0.1	-0.3	-0.2
Jes eff. mix. 1	0.1	-0.1	-0.1	-0.1	-1.2	0.6	-0.4	0.4	-0.2	-0.1	0.1	-0.1
Jes eff. mix. 2	0.0	-0.0	-0.0	-0.3	1.8	-1.1	0.6	-0.5	0.1	-0.4	-0.4	-0.1
Jes eff. mix. 3	-0.0	0.1	-0.2	-0.0	-0.2	0.9	-0.1	0.2	-0.7	0.1	-0.3	-0.4
Jes eff. mix. 4	0.0	0.0	-0.1	-0.1	0.0	-0.0	0.1	0.1	-0.5	-0.6	-0.7	-0.1
Jes eff. model 1	-0.1	-0.2	-1.4	0.2	-3.2	3.6	-3.2	3.4	1.1	-1.3	0.6	-2.1
Jes eff. model 2	0.0	0.0	-0.1	-0.1	0.8	-0.4	0.1	-0.0	0.0	-0.4	-0.4	-0.7
Jes eff. model 3	-0.1	0.1	-0.3	0.1	-1.0	1.1	1.4	0.3	-0.9	0.2	-0.7	-0.1
Jes eff. model 4	0.0	-0.0	-0.1	-0.2	-0.0	-0.2	0.0	0.1	-0.5	-0.6	-0.4	-0.5
Jes eff. stat. 1	-0.1	0.1	-0.7	0.2	-2.2	1.6	-1.3	1.4	0.0	-0.5	-0.2	-0.3
Jes eff. stat. 2	0.0	-0.0	-0.1	-0.2	0.6	-0.0	0.1	-0.1	-0.1	-0.7	-0.2	-0.5
Jes eff. stat. 3	-0.2	0.2	-0.3	0.1	0.3	-0.0	-0.0	0.1	-1.0	-0.3	-0.4	0.2
Jes eff. stat. 4	0.0	0.0	-0.1	-0.2	-0.3	1.0	-0.1	0.2	-0.4	-0.2	-0.3	-0.3
Jes pile-up μ	0.0	-0.1	0.4	-1.0	1.8	0.2	0.6	-0.5	-0.3	-0.1	-0.3	-0.1
Jes pile-up # PV	-0.1	-0.0	-0.2	-0.5	1.0	-0.6	0.8	-0.7	-0.5	0.3	-0.4	1.0
Jes pile-up p_{T}	0.0	-0.0	-0.1	-0.2	-0.3	-0.2	0.0	0.0	-0.5	-0.6	-0.1	-0.2
Jes pile-up p_{T} dens.	-0.2	0.0	-1.2	0.3	-2.8	3.1	-2.5	2.6	0.7	-0.9	0.1	-0.9
Jes Flavour comp.	-1.8	1.7	-3.9	2.9	-4.7	3.8	-5.0	5.3	-1.6	1.2	-2.2	0.5
Jes Flavour resp.	1.1	-1.2	2.1	-2.8	2.5	-3.8	3.6	-3.3	0.5	-0.5	0.0	-2.1
b-Jes	1.1	-1.1	1.3	-1.6	0.4	-0.7	0.7	-0.6	0.6	-2.1	1.3	-1.5
Jes Punch through	0.0	0.0	-0.2	-0.2	-0.3	-0.3	0.1	0.1	-0.5	-0.6	-0.5	-0.6
Jet energy res.	-3.2	3.2	-2.8	2.8	-3.0	3.0	-2.1	2.1	1.2	-1.2	-0.6	0.6
Jet reco. eff.	0.0	-0.0	0.0	-0.0	0.5	-0.5	0.1	-0.1	0.1	-0.1	0.1	-0.1
Jet vertex fraction	0.1	-0.2	0.1	-0.2	1.4	-0.6	1.1	-0.9	0.2	0.0	0.2	-0.4
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft res.	0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.0	0.0	0.0	0.0	0.4	0.6
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft scale	0.1	-0.2	0.1	-0.2	-0.4	-0.0	0.0	-0.1	0.6	-0.1	0.9	-0.9
Lepton reco.	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2
Lepton Ident.	1.3	-1.3	1.3	-1.3	1.3	-1.3	1.3	-1.3	1.2	-1.2	1.2	-1.2
Lepton trigger	-0.8	0.8	-0.8	0.8	-0.8	0.8	-0.8	0.8	-0.9	0.9	-0.9	0.9
El. energy res.	0.1	-0.0	-0.3	-0.1	1.0	0.0	0.0	0.0	-0.4	-0.1	-0.5	-0.2
El. energy scale.	0.8	-0.8	0.6	-0.8	0.9	0.4	0.4	-0.4	0.6	-0.6	0.3	-0.3
Muon ID mom. res.	0.0	-0.0	0.1	-0.1	-0.0	0.0	-0.0	0.0	0.0	-0.0	-0.0	0.0
Muon MS mom. res.	0.0	-0.0	-0.0	0.0	-0.5	0.5	-0.0	0.0	0.0	-0.0	0.3	-0.3
Muon mom. scale	0.0	-0.1	0.1	-0.0	0.3	0.1	0.0	-0.0	-0.0	-0.0	0.0	0.0
b-tagging	4.6	-4.5	3.9	-3.9	3.1	-3.1	4.3	-4.2	2.5	-2.4	3.5	-3.4
c-tagging	0.0	-0.0	1.2	-1.2	3.3	-3.3	0.6	-0.6	4.0	-3.9	1.5	-1.4
mis-tagging	0.1	-0.1	1.5	-1.5	1.5	-1.5	0.2	-0.2	7.4	-6.6	3.3	-3.0
PDFs	1.3	-1.3	2.0	-2.0	2.3	-2.3	3.1	-3.1	2.9	-2.9	3.5	-3.5
s-channel gen.	-0.2	0.2	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0
s-channel gen. scales	6.6	-7.3	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0
t-channel gen.	0.0	-0.0	8.8	-8.8	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0
Wt gen.	0.0	-0.0	0.0	-0.0	-8.2	8.2	0.0	-0.0	0.0	-0.0	0.0	-0.0
$t\bar{t}$ gen.	0.0	-0.0	0.0	-0.0	0.0	-0.0	-3.0	3.0	0.0	-0.0	0.0	-0.0
$t\bar{t}$ modelling	0.0	-0.0	0.0	-0.0	0.0	-0.0	2.7	-6.7	0.0	-0.0	0.0	-0.0
W+jets gen.	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0	-9.7	9.7	0.0	-0.0

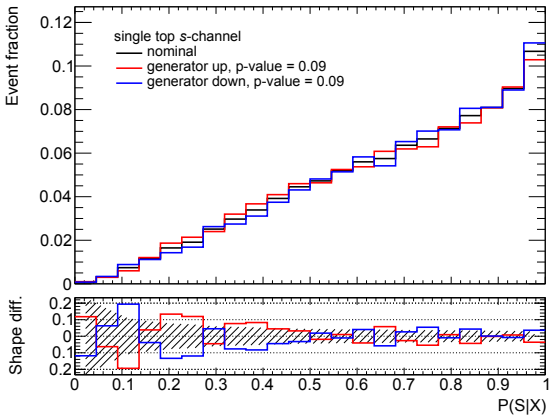
Table 3.8.: Rate uncertainties in the signal region. Only a few of the sources of uncertainty have a significant impact on the signal and background event rates. In particular these are the normalisations of background processes based on theoretical predictions, modelling uncertainties estimated by comparisons of predictions given by different MC event generators and the b-tagging.



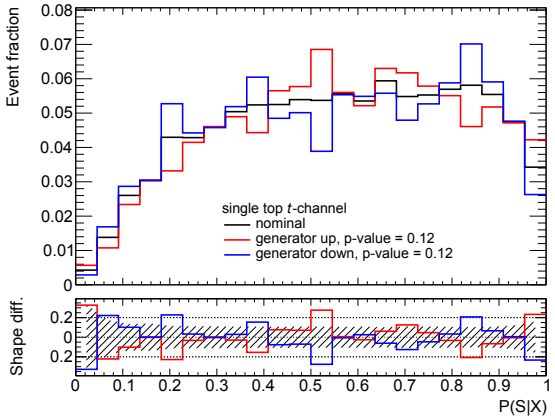
(a) Jet energy resolution, s-channel



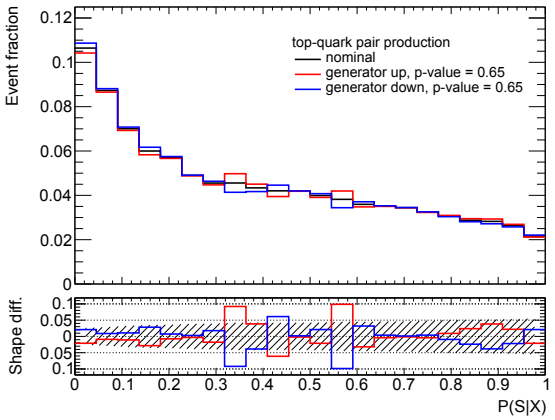
(b) Jet energy resolution, t-channel



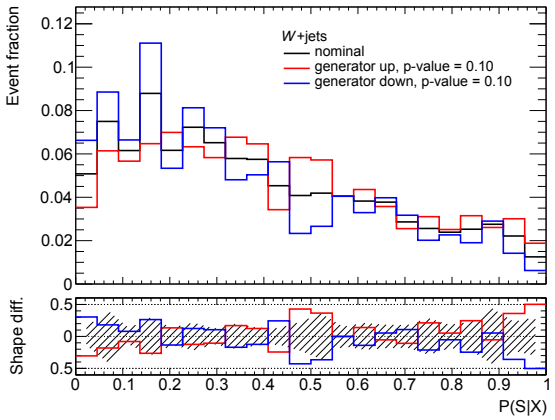
(c) s-channel generator



(d) t-channel generator



(e) $t\bar{t}$ generator



(f) W+jets generator

Figure 3.20.: Examples of systematic variations of the shape of the TELL discriminant in the signal region. The uncertainty bands shown in the lower panels present statistical uncertainties. In (a) and (b) the statistical uncertainty of the nominal distribution is shown, whereas in the other cases the larger statistical uncertainties related to the alternative event simulation are shown.

systematically varied distributions. By using disjoint event sets, correlations between the pairs of distributions entering the test are avoided. In consequence, the results of the tests can be properly interpreted in terms of their p-values. If the pairs of distributions were not derived from disjoint event sets, the resulting χ^2-values could not easily be converted into p-values. An exception to this relation is given by the case of modelling uncertainties, where the set of systematically varied events is generated differently. Hence there are no correlations between the nominal and the systematically varied distributions. Accordingly, events are not split into subsets in this case but the complete sets are used instead.

Figure 3.20 shows examples of comparisons of shapes of the TELL discriminant $P(S|X)$, together with the resulting p-values. For each systematic uncertainty the χ^2-tests are performed for the upward and for the downward variation. In the following, the smaller p-value is used. In case of shape variations which are smaller than the statistical uncertainty of the simulation, the p-value is equally distributed between zero and one. On the other hand, in case of significant shape variations small p-values are expected. In the context of the statistical model discussed below in Sec. 3.4, systematic uncertainties are included as shape variations if the corresponding p-value is less than 5%, while all other systematic variations remain as overall rate uncertainties.

This choice defines a reasonable but somewhat arbitrary criterion of a significant shape difference. In App. C.5 it is shown that it only has a marginal impact on the measurement except for the treatment of exactly one source of uncertainty, namely the modelling of $W+$jets events, where the Sherpa generator is compared with Alpgen interfaced with Pythia 6. Since such comparisons of event generators can hardly be interpreted in terms of discriminant shape interpolations if their impact on the result is significant, a dedicated study of this uncertainty is presented in App. C.5. The main result of this study is the observation that the signal cross section determined by a fit of the model to the data depends on the choice of the $W+$jets event generator which is used by this model. This observed deviation matches the expectation based on a comparison of the two corresponding models. Furthermore, the difference of the discriminant shape based on the two predictions is insignificant as shown in Fig. 3.20f. The p-value of the corresponding χ^2-test amounts to 10%. Furthermore, the studies of kinematic distributions as well as TELL distributions presented above show that the baseline $W+$jets event generator Sherpa, which comes with more generated events than the Alpgen + Pythia 6 sample, provides a proper modelling of the data. Hence the Sherpa generator is chosen to model $W+$jets events while the overall rate uncertainty

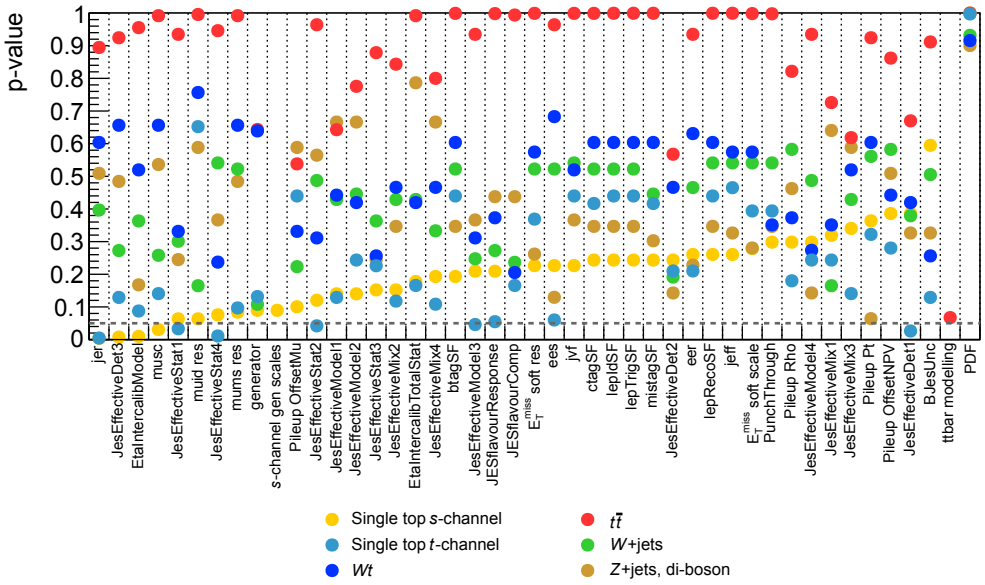


Figure 3.21.: Discriminant shape variations due to systematic uncertainties. The p-values of χ^2-tests comparing nominal and systematically varied discriminant distributions are shown for each simulated process. The goodness-of-fit test is performed for the upward and downward variations independently and the lower p-value is chosen. Uncertainties giving p-values below 5% are included as shape variations in the statistical model discussed in Sec. 3.4.

based on the W+jets generator comparison is taken into account. A more precise comparison of these generators demands for increased simulation statistics. Thus it remains to be carried out in the context of future measurements.

All p-values resulting from the χ^2-tests are shown in Fig. 3.21 for all of the simulated processes. Almost all χ^2-tests lead to intermediate p-values, as is expected in case of insignificant shape variations. Rather high p-values are obtained in case of the PDF variations. These p-values are an exception. They are not necessarily distributed equally between zero and one, because here by construction the varied distributions are correlated with the nominal ones to which they are compared according to Eq. 3.35. A few uncertainties lead to small p-values below 5%. Table 3.9 summarises these uncertainties, for which the shape variation is considered to be significant.

3.4. Search for s-channel single top-quark production

Based on the preparations discussed in the previous sections, the number of s-channel single top-quark events contained in the given data set can be measured. In order to

Systematic uncertainty	*s*-channel	*t*-channel	all other processes
η-intercalibration modelling	✓	–	–
JesEffectiveDet1	–	✓	–
JesEffectiveDet3	✓	–	–
JesEffectiveModel3	–	✓	–
JesEffectiveStat1	–	✓	–
JesEffectiveStat2	–	✓	–
JesEffectiveStat4	–	✓	–
jet energy resolution	✓	✓	–
muon momentum scale	✓	–	–

Table 3.9.: Summary of discriminant shape variations due to systematic uncertainties. The "JesEffective" uncertainties are associated with the jet energy scale and result from nuisance parameter combinations. Only systematic uncertainties which lead to significant shape variations at least for one scattering process are shown. Such shape variations are marked. They are considered in the statistical model discussed in Sec. 3.4.

carry out this measurement, a statistical model of the data must be constructed on the basis of the signal and background predictions. More precisely, the model is a probability density function (Pdf) of the TELL discriminant $P(S|X)$. This observable is suited for the identification of a possible *s*-channel signal. The model must take statistical and systematic uncertainties into account and it necessarily contains several parameters. On the one hand, there are the nuisance parameters which are related to the systematic uncertainties. They are constrained by the auxiliary knowledge discussed in the previous section. On the other hand, there is the frequency of signal events, or the signal strength, which is the parameter of interest. It is defined by the ratio of the measured signal cross section over the corresponding SM prediction,

$$\mu = \frac{\sigma^{\text{observed}}_{s\text{-channel}}}{\sigma^{\text{SM}}_{s\text{-channel}}} \ . \tag{3.36}$$

Apart from the obvious limitation $\mu \geq 0$ the model does not constrain it in any way. The Pdf of $P(S|X)$ can also be seen as a function of its parameters, given the measured data. The resulting likelihood function is suited to fit the parameters to the data. This procedure constitutes the measurement. In this context, the first question which needs to be answered is whether *s*-channel single top-quark events are observed at a significant rate. Further studies depend on the answer to this question. In the following, the

likelihood function is introduced first. Then the data is interpreted in view of s-channel single top-quark production.

3.4.1. Profile likelihood fit

The statistical model makes use of histograms of the distribution of $P(S|X)$ as predicted by the employed simulations and lepton fake estimates. These histograms are combined into a Pdf of the discriminant using the computer program HistFactory [195]. The statistical analysis itself is based on the RooStats [196], RooFit [197] and ROOT [188] programs. One part of the model is a product of single event probability densities of the discriminant. For a single process, this Pdf reads

$$f_{\text{proc}}(x_e) = \frac{1}{\Delta_b} \frac{\nu_{b_e}^{\text{proc}}}{N_{\text{proc}}} \ , \tag{3.37}$$

where the following notations are used:

- x_e, b_e: The discriminant $P(S|X)$ measured in the event e is simply denoted by x_e. b_e is the index of the corresponding bin of the discriminant histogram;
- Δ_b : The width of the histogram bins. An equidistant binning is used;
- $\nu_{b_e}^{\text{proc}}$: The number of events of the process "proc" in the bin b_e;
- N_{proc}: The total number of events of the process "proc".

The overall Pdf of x_e is given by the sum of the individual contributions, weighted by their event fractions,

$$\frac{\mu S f_S(x_e) + \sum_{k\in\{\text{BKGs}\}} B_k f_{\text{bkg k}}(x_e)}{\mu S + B} = \frac{1}{\Delta_b} \frac{\mu \nu_{b_e}^{\text{sig}} + \sum_{k\in\{\text{BKGs}\}} \nu_{b_e}^{\text{bkg k}}}{\mu S + B} \ . \tag{3.38}$$

The value S gives the total number of signal events according to the SM prediction $\mu = 1$. B_k denotes the total number of events of the background k and B denotes the sum of these events,

$$S = \sum_{b\in\{\text{bins}\}} \nu_b^{\text{sig}}, \qquad B_k = \sum_{b\in\{\text{bins}\}} \nu_b^{\text{bkg k}}, \qquad B = \sum_{k\in\{\text{BKGs}\}} B_k \ . \tag{3.39}$$

The model of the overall measurement of the events $\{x_e\}$ consists of three components. Firstly, the product of single event Pdfs given by Eq. 3.38 over all events gives the total Pdf for a given number of events. The bin width Δb can be neglected because it is constant. Thus it does not matter in the context of maximum likelihood fits. Secondly, the total number of events is subject to the measurement as well. A Poissonian term enters the model accordingly. In addition, the number of events of each process in each bin, $\nu_{b_e}^{\text{proc}}$, is subject to systematic uncertainties which are introduced by means of the nuisance parameters, which in turn are subject to auxiliary constraints. These constraints enter the model as well. Altogether, the likelihood function reads

$$L(n, \boldsymbol{x_e}|\mu, \alpha) = \text{Pois}(n|\mu S + B) \prod_{x_e \in \{\text{events}\}} \frac{\mu \nu_{b_e}^{\text{sig}} + \sum_{k \in \{\text{BKGs}\}} \nu_{b_e}^{\text{bkg k}}}{\mu S + B} \prod_{p \in \{\text{syst}\}} G(0|\alpha_p, 1) \ . \tag{3.40}$$

The number of events of each process in each bin, $\nu_{b_e}^{\text{proc}}$, depends on the corresponding nominal and systematically varied histograms and on the nuisance parameters,

$$\nu_{b_e}^{\text{proc}} = \lambda^{\text{proc}} \gamma_{b_e} \eta_{\text{proc}}(\alpha) \sigma^{\text{proc}}(x_{b_e}, \alpha) \ , \tag{3.41}$$

$$\eta_{\text{proc}}(\alpha) = \prod_{p \in \{\text{syst}\}} \left(\eta_{\text{proc}}^{p\,\pm}\right)^{|\alpha_p|} \ , \tag{3.42}$$

$$\sigma_{\text{proc}}(x_{b_e}, \alpha) = \sigma_{\text{proc}}^{\text{nom}}(x_{b_e}) + \sum_{p \in \{\text{syst}\}} |\alpha_p| \left(\sigma_{\text{proc}}^{p\,\pm}(x_{b_e}) - \sigma_{\text{proc}}^{\text{nom}}(x_{b_e})\right) \ . \tag{3.43}$$

The following notations are used:

- n: The total number of measured events. It is distributed according to $\text{Pois}(n|\mu S + B) = (\mu S + B)^n \, e^{-(\mu S + B)}/n!$;
- λ^{proc}: Integrated luminosity. The nominal value $L = 20.3\,\text{fb}^{-1}$ corresponds to $\lambda = 1$. There is exactly one luminosity nuisance parameter λ for all processes which are estimated by means of simulations. In case of the lepton fake contribution this nuisance parameter does not enter since this is a data-driven estimation;
- γ_{b_e}: Nuisance parameter of the statistical uncertainty of the predictions in the bin b_e. There is one such nuisance parameter per bin. This is a simplified version of the Barlow–Beeston approach [198];
- $G(0|\alpha_p, 1)$: Constraint term of the systematic uncertainty with index p. The corresponding nuisance parameter is denoted by α_p. Standard Gaussian distributions

are used, $G(0|\alpha_p, 1) = e^{-\alpha_p^2/2}/\sqrt{2\pi}$. The set of all nuisance parameters α also includes the luminosity nuisance parameter λ and the statistical uncertainty nuisance parameters γ_{b_e}. In theses cases the widths of the constraint terms are adjusted accordingly;

- $\eta_{\text{proc}}(\alpha)$: A factor which combines the effect of all normalisation uncertainties apart from the luminosity uncertainty. There is one such factor per process. $\eta_{\text{proc}}^{p\,\pm}$ denotes the relative deviation of the number of events of the process at hand with regard to the systematic variation p in the positive and negative direction, respectively, by one standard deviation. The deviations from the nominal case $\eta = 1$ are parametrised by means of the nuisance parameters α_p using an exponentiation. Together with the employed Gaussian constraint terms G, this procedure is equivalent to the usage of a linear interpolation and log-normal constraint terms. The choice between η_{proc}^{p+} and η_{proc}^{p-} depends on the sign of α_p;

- $\sigma_{\text{proc}}(x_{b_e}, \alpha)$: The histogram of the distribution of the observable given the process at hand, evaluated at the bin b_e. The nominal histogram is denoted by $\sigma_{\text{proc}}^{\text{nom}}$. In case of systematic variations of the shape of the distribution, the second term on the right-hand side of Eq. 3.43 enters. $\sigma_{\text{proc}}^{p\,\pm}$ denotes the discriminant distribution given a variation of the systematic uncertainty p in the positive or negative direction, respectively, by one standard deviation. The nuisance parameters α_p are used to linearly parametrise deviations from the nominal case. The choice between $\sigma_{\text{proc}}^{p+}$ and $\sigma_{\text{proc}}^{p-}$ depends on the sign of α_p.

The likelihood function given by Eq. 3.40 is evaluated using the n measured events x_e. Thus it is a function of the signal strength μ and of the nuisance parameters α. Accordingly, it is simply denoted by $L(\mu|\alpha)$ in the following.

The measurement makes use of maximum likelihood fits of the parameters given the measured data. For this purpose the following quantity is used,

$$q_\mu = -2 \ln \frac{L(\mu|\hat{\hat{\alpha}})}{L(\hat{\mu}|\hat{\alpha})} \; . \tag{3.44}$$

The denominator $L(\hat{\mu}|\hat{\alpha})$ denotes the global maximum of the likelihood function. In other words, $\hat{\mu}$ and $\hat{\alpha}$ are the unconditional maximum likelihood estimators (MLEs) of the parameters μ and α, respectively. The nominator $L(\mu|\hat{\hat{\alpha}})$, on the other hand, is the maximum of the likelihood function given the signal strength μ. Thus it is a one-dimensional function of the latter and $\hat{\hat{\alpha}}$ denotes the conditional MLEs of the nuisance

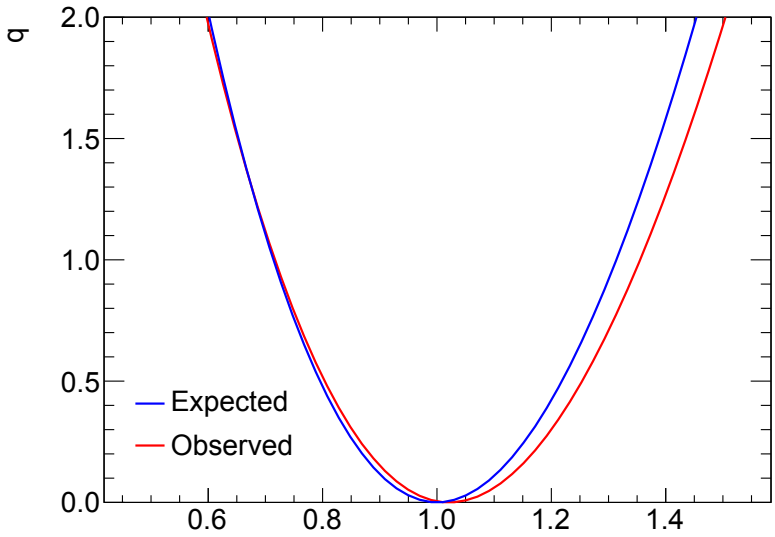


Figure 3.22.: Profile likelihood curve according to Eq. 3.44. The likelihood function is fit to the Asimov data set, shown in blue, and to the measured data set, shown in red. The observed curve is close to the expectation.

parameters α. In consequence, q_μ is a non-negative, one-dimensional function of μ. The lower its value the better the fit to the data. For the MLE $\mu = \hat{\mu}$ one obtains $q_{\hat{\mu}} = 0$. The width of the q_μ-curve is related to the precision of the measurement of μ. The nuisance parameters α, representing the systematic uncertainties, lead to a broadening of this curve. The conditional maximisation of the likelihood function, $L(\mu|\hat{\hat{\alpha}})$, allows one to gather information on these uncertainties in addition to the measurement of the signal strength. This technique is know as profiling and q_μ is called the profile likelihood ratio.

In order to search for *s*-channel single top-quark production, the profile of the likelihood function is determined. First of all, the procedure is performed on the basis of the expected data given the SM hypothesis and the auxiliary constraints on systematic uncertainties. More precisely, the Asimov data set is used [199]. In every bin of the histogram of the discriminant distribution the Asimov data is defined by the event yield which is predicted by the model. Hence this artificial data set would be obtained in a measurement if the model was perfectly true and if it wasn't for statistical fluctuations of the data itself. The likelihood curve which results from fits to the Asimov data set is shown in Fig. 3.22. The maximum of the likelihood function is reached at the SM value $\mu = 1$. Thus the fit is unbiased. Finally, the procedure is performed on the basis of the measured data. The result is shown in Fig. 3.22 as well. The shape of the observed curve is parabolic, which matches the expectation, while its width is slightly enhanced. The

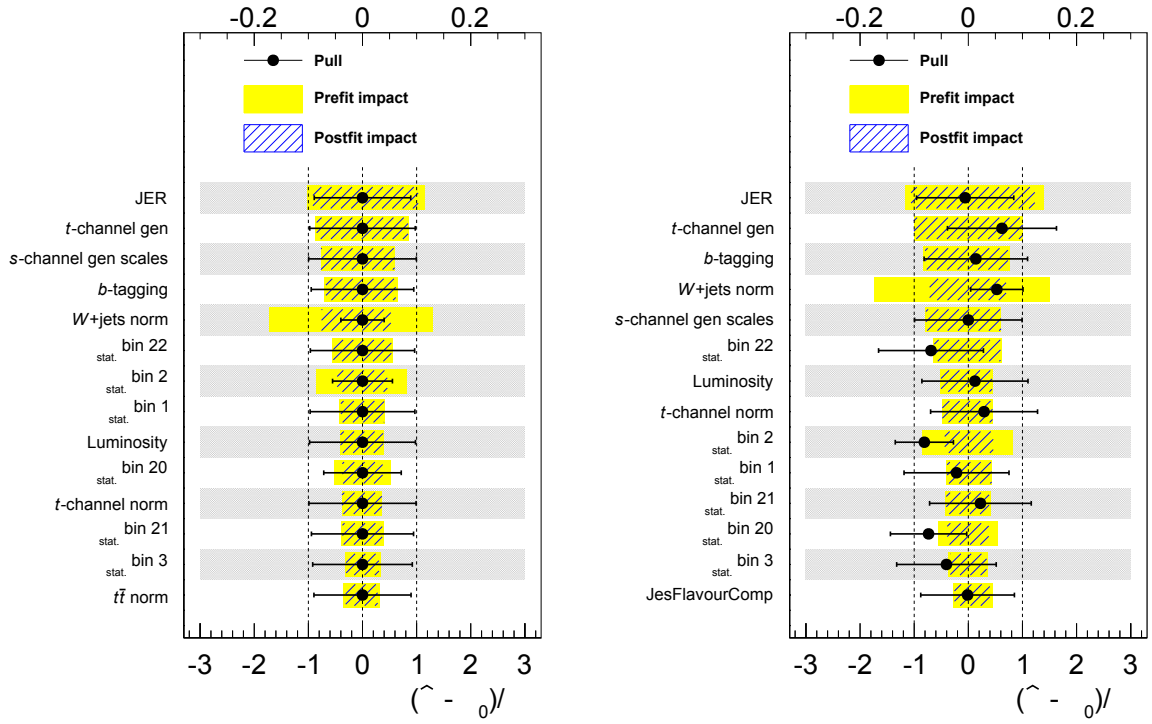
observed MLE of the signal strength is

$$\hat{\mu}_{s\text{-channel}} = 1.02 \ . \tag{3.45}$$

Hence the observed result is very close to the SM expectation.

A closer look at the fit is taken in Fig. 3.23, where 3.23a refers to the Asimov data and 3.23b refers to the measured data. Firstly, both figures show the MLEs of nuisance parameters α, normalised with respect to the corresponding input values and uncertainties. The uncertainties obtained from the fit are provided as well. For their estimation the matrix of second order derivatives of the likelihood function is used. Most of the MLEs and post-fit uncertainties are close to the expectation α_0 and $\Delta\alpha$, respectively. There are a few nuisance parameters which are constrained considerably by the fit. In particular, this is true for the W+jets normalisation uncertainty. Its input uncertainty is large. Therefore it is expected to be constrained by the fit. This is shown explicitly by the fit to the Asimov data. Furthermore, nuisance parameters corresponding to the statistical uncertainties of the predictions are also constrained by the fit. The discriminant distribution given by the model is subject to fluctuations caused by low statistics of predictions, in particular of the W+light-flavour jets estimate. The fit effectively smoothes this distribution, which results in constraints on the aforementioned nuisance parameters. While Fig. 3.23 only involves the most important nuisance parameters, App. C.6 provides figures which show the remaining ones. Overall, a reasonable fit result is observed.

Secondly, Fig. 3.23 also shows the impact of nuisance parameter variations on the signal strength. The fit is repeated while the nuisance parameter in question is varied by one standard deviation and fixed accordingly. The resulting deviation of the MLE of the signal strength from the result of the full fit is shown for each nuisance parameter. This procedure is performed on the basis of the model parameters given before and after the fit, respectively. As a result, the jet energy resolution uncertainty turns out to have the largest impact on the signal strength. This observation is in accordance with the fact that jets play a key role in the analysis, apart from electrons and muons, which are reconstructed with a higher precision. The second largest impact on the signal strength is related to the modelling of t-channel single top-quark production background events. The corresponding MC event generator variation introduces a relatively large rate uncertainty of 9%. Given the similarity of the s and t-channel discriminant distributions, this variation affects the signal strength considerably. Also the b-tagging uncertainty has a large impact because it affects the rate of all simulated scattering processes to a large

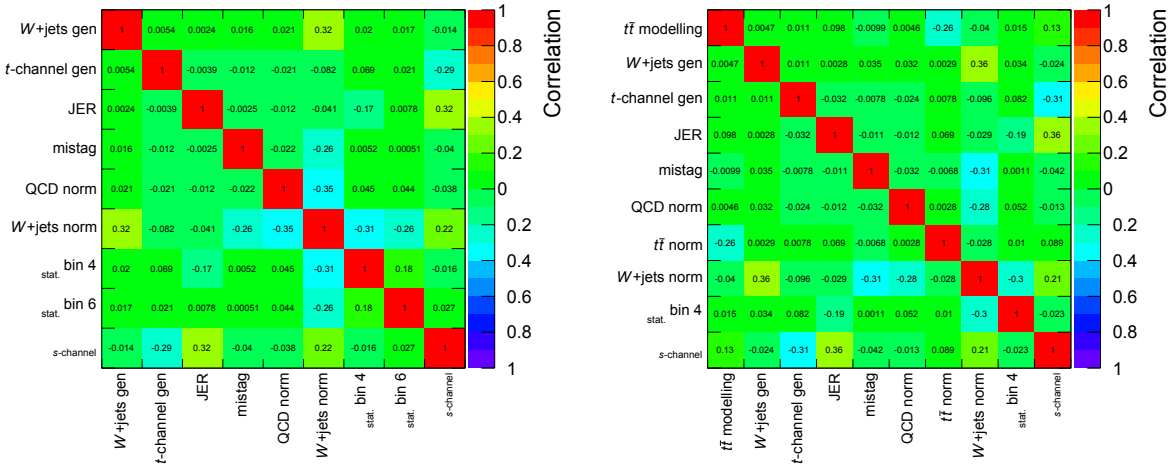


(a) Expectation **(b)** Measurement

Figure 3.23.: Nuisance parameter fit results based on the Asimov data (a) and on the measured data (b). The pulls provide the MLEs of nuisance parameters together with their fit uncertainties. Only a few nuisance parameters are constrained by the fit. In addition, the impact of nuisance parameters on the signal strength is provided. The observed impacts are very similar to the expectation.

extent. Furthermore, the normalisation of the W+jets background has a large impact too, because this process is the second most frequent background and its normalisation uncertainty is large. Furthermore, there are many nuisance parameters related to MC statistical uncertainties which have a large impact. Overall it must again be noted that the observed fit results are very similar to the expected ones.

Figure 3.24 shows correlations between the fit parameters. Again the matrix of second order derivatives of the likelihood function is used. Weakly correlated nuisance parameters are neglected in this figure. Again the measured results are very similar to the expectation. In particular, the signal strength is correlated most strongly with the jet energy resolution, the t-channel modelling and the W+jets normalisation, which is similar to the case of the impacts on the signal strength discussed above. Another observation is the fact that the W+jets normalisation uncertainty is strongly correlated



(a) Expectation **(b)** Measurement

Figure 3.24.: Fit parameter correlations based on the Asimov data (a) and on the observed data (b). The two results are very similar to each other. Only parameters which have a correlation of more than 25% with at least one other parameter are shown.

with several other uncertainties. This is expected due to the largeness of this uncertainty. Also the simulated W+jets sample involves only few events, so that MC statistical uncertainties and the overall W+jets normalisation are correlated with each other.

This study of the maximum likelihood fit to the observed data and its comparison with the expectation shows that the employed model of the data is valid. Finally, Fig. 3.25 shows the discriminant distribution according to the model where the fit result is taken into account. In addition, the measured distribution is shown. Indeed the model fits the data. Concerning the s-channel signal, the figure suggests that the model must involve this process in order to allow for a proper fit to the data. This observation is emphasized by Fig. 3.26, which is derived from the previous figure, subtracting the number of background events as predicted by the model after the fit in each bin.

3.4.2. Signal significance

So far the discussion of the maximum likelihood fits shows that the model of the discriminant distribution provides an appropriate description of the measured data. In particular, the s-channel signal is a vital part of the model. Now this observation needs to be turned into a quantitative result. For this purpose the level of incompatibility of

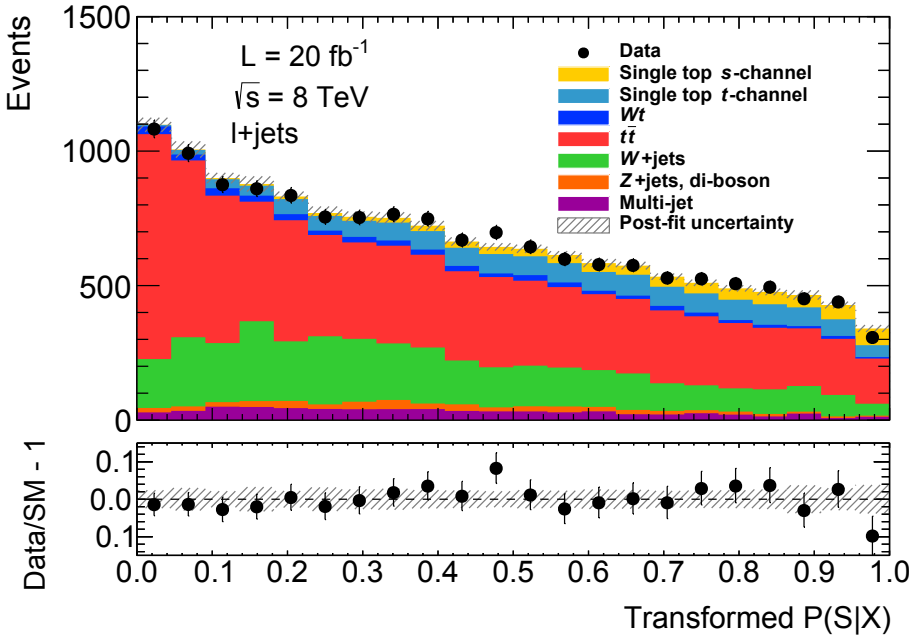


Figure 3.25.: Post-fit discriminant distribution. The model fits the data well. The *s*-channel single top-quark production process is an important part of the model.

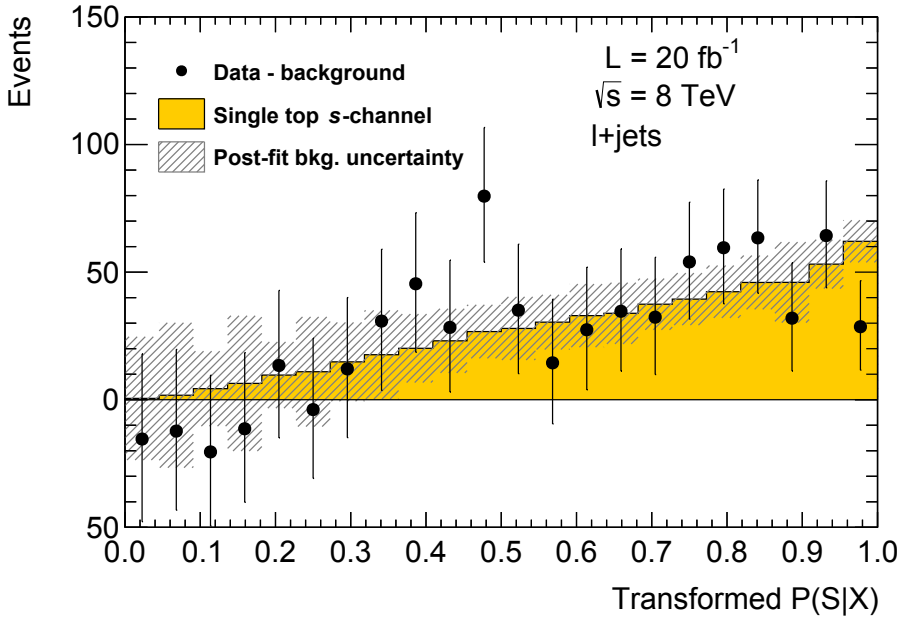


Figure 3.26.: Post-fit discriminant distribution with background subtraction. The uncertainty band only takes into account the nuisance parameters. The complete distribution shown in Fig. 3.25 is adjusted by a subtraction of the background events according to the model after the fit. The remaining *s*-channel signal process is needed in order to explain the measured data.

the background-only hypothesis with the measured data is determined. More precisely, the simple hypothesis $\mu = 0$ is tested against the signal hypothesis $\mu > 0$. The employed test statistic is

$$\tilde{q}_0 = \begin{cases} q_{\mu=0} = -2\ln\left(\frac{L(0|\hat{\hat{\alpha}}(0))}{L(\hat{\mu}|\hat{\alpha})}\right) , & \text{if} \quad \hat{\mu} \geq 0 \\ 0 , & \text{if} \quad \hat{\mu} < 0 \end{cases} . \tag{3.46}$$

It coincides with the profile likelihood ratio q_μ with $\mu = 0$ in case of non-negative MLEs of the signal strength. Small values of $\tilde{q}_0$ support the background-only hypothesis because here typically $\hat{\mu} \approx 0$. Large values, on the other hand, support the signal hypothesis. Negative values of $\hat{\mu}$ do not indicate a deviation from the background-only hypothesis. Therefore the value zero is assigned to the test statistic in this case. In order to quantify the level of disagreement between the data and the background-only hypothesis, the following p-value is used, [2]

$$p_0 = \int_{\tilde{q}_0^{\text{obs}}}^{\infty} f(\tilde{q}_0|\,\mu = 0)\,\mathrm{d}\tilde{q}_0 , \tag{3.47}$$

where $f(\tilde{q}_0|\,\mu{=}0)$ denotes the Pdf of $\tilde{q}_0$ under the background-only hypothesis $\mu = 0$ and $\tilde{q}_0^{\text{obs}}$ denotes the observed value of the test statistic. The Pdf is integrated up to infinity because large values of $\tilde{q}_0$ support the signal hypothesis, against which the background hypothesis is tested.

In order to determine the Pdf $f(\tilde{q}_0|\,\mu{=}0)$, pseudo-experiments are performed where variations due to statistical fluctuations as well as systematic uncertainties are taken into account by means of an MC technique. In each of the pseudo-experiments, first of all the argument α_{obs} of each constrain term G is thrown randomly according to this standard Gaussian distribution. Hence $G(0|\alpha, \sigma_\alpha{=}1)$ becomes $G(\alpha_{\text{obs}}|\alpha, \sigma_\alpha{=}1)$. The argument α_{obs} is fixed after this step. In a second step, the resulting model is used to generate a distribution of the discriminant. Then the test statistic $\tilde{q}_0$ is determined on the basis of this pseudo-data set. For this purpose two maximum likelihood fits are performed, giving the nominator and the denominator of Eq. 3.46. The whole procedure is repeated in order to fill a histogram of $\tilde{q}_0$ values which provides the desired Pdf $f(\tilde{q}_0|\,\mu{=}0)$. Concerning the background-only hypothesis, 1.6 million pseudo-experiments are performed, so that

[2] The p-values are converted into significances based on standard Gaussian distributions. For a given one-sided p-value, the distance between the corresponding argument and the mean value zero defines the significance.

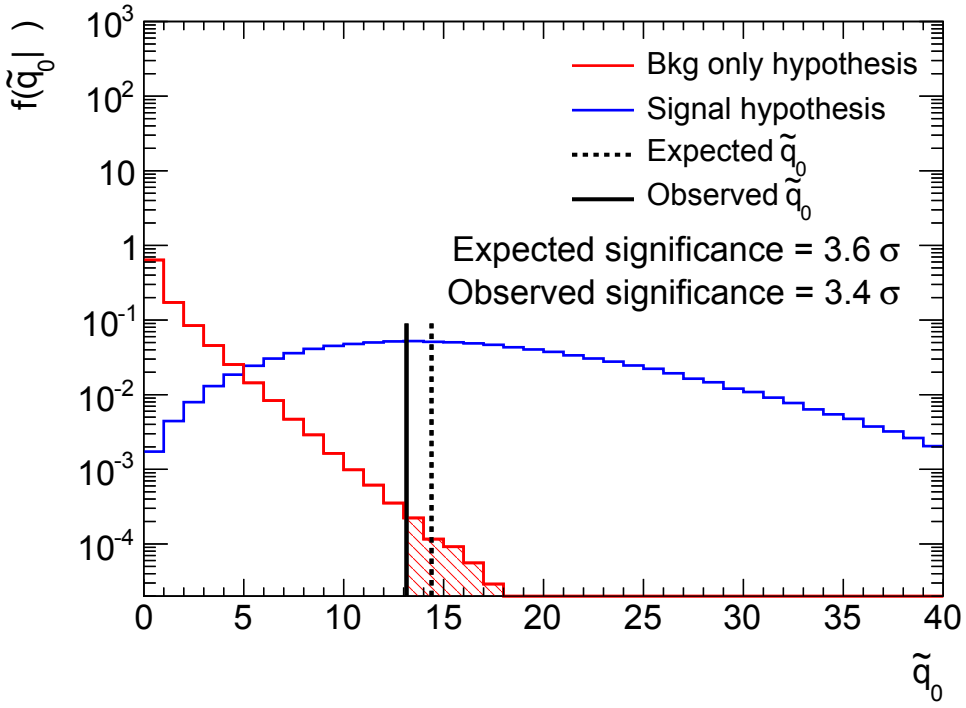


Figure 3.27.: Search for s-channel single top-quark production. Pseudo-data is generated based on the background-only hypothesis $\mu = 0$ as well as on the signal hypothesis $\mu = 1$ in order to determine the corresponding distributions of the test statistic $\tilde{q}_0$. The observed test statistic can hardly be explained by means of the background-only hypothesis but it is consistent with the signal hypothesis. This is evidence for s-channel single top-quark production. The result is close to the SM expectation.

the statistical uncertainties of the desired p-values are small. Finally, $\tilde{q}_0^{\text{obs}}$ is determined using the Asimov data and the measured data, respectively, so that the expected and observed p-values p_0 can be derived.

The result of this approach is presented in Fig. 3.27. Apart from the background-only distribution $f(\tilde{q}_0|\,\mu = 0)$, also the distribution of the signal hypothesis $f(\tilde{q}_0|\,\mu = 1)$ is shown. It is derived from pseudo-experiments as well. The two distributions clearly differ from each other. In consequence, the test statistic based on the Asimov data set leads to a small p-value of $p_0^{\text{expected}} = (1.6 \pm 0.1) \cdot 10^{-4}$, or, equivalently, to a significance of 3.60 ± 0.02 standard deviations. This result proves the ability of the analysis to separate the s-channel signal from the backgrounds. Finally, the test statistic is determined on

the basis of the measured data. This results in an observed p-value of

$$\begin{aligned} p_0^{\text{observed}} &= (2.9 \pm 0.1) \cdot 10^{-4}, \\ &\Leftrightarrow \\ \text{significance} &= 3.45 \pm 0.01 \text{ standard deviations} . \end{aligned} \tag{3.48}$$

This is the first evidence for s-channel single top-quark production in pp collisions. Again, the observed result is close to the SM expectation.

3.4.3. Cross section confidence interval

In order to estimate the constraints on the signal cross section given by the fit of the model to the data, a confidence interval is constructed. This task can be accomplished on the basis of the likelihood curve shown in Fig. 3.22. However, a more general approach is chosen in the following, where the signal hypothesis is tested against the background-only hypothesis. This is done for a whole set of signal hypotheses, scanning the possible values of the signal strength $\mu > 0$. The resulting confidence interval turns out to equal the one derived from the likelihood curve. In addition, the chosen approach also allows one to underline the evidence for the s-channel signal by using pseudo-data distributed according to the background-only hypothesis.

For the hypothesis tests a symmetric test statistic is used, namely q_μ as already provided by Eq. 3.44. The signal is now known to exist up to a high significance. Furthermore, the signal hypothesis according to which the data is assumed to be distributed is defined by the MLE of the signal strength, $\mu^{\text{signal}} = \hat{\mu}$. The test statistic q_μ is small within a range of signal strengths $\mu \approx \hat{\mu}$. In case of data distributed according to the background-only hypothesis, on the other hand, the MLE $\hat{\mu}$ is close to zero. Hence for every tested signal strength $\mu > 0$ the test statistic is relatively large in this case. Accordingly, every hypothesis $\mu > 0$ is tested against the background-only hypothesis by means of the following p-value,

$$p_\mu = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu |\, \mu^{\text{signal}})\, \mathrm{d}q_\mu \ , \tag{3.49}$$

where $f(q_\mu |\, \mu^{\text{signal}})$ denotes the distribution of q_μ given the signal hypothesis. For each test it is determined by means of pseudo-experiments. They are carried out in the same

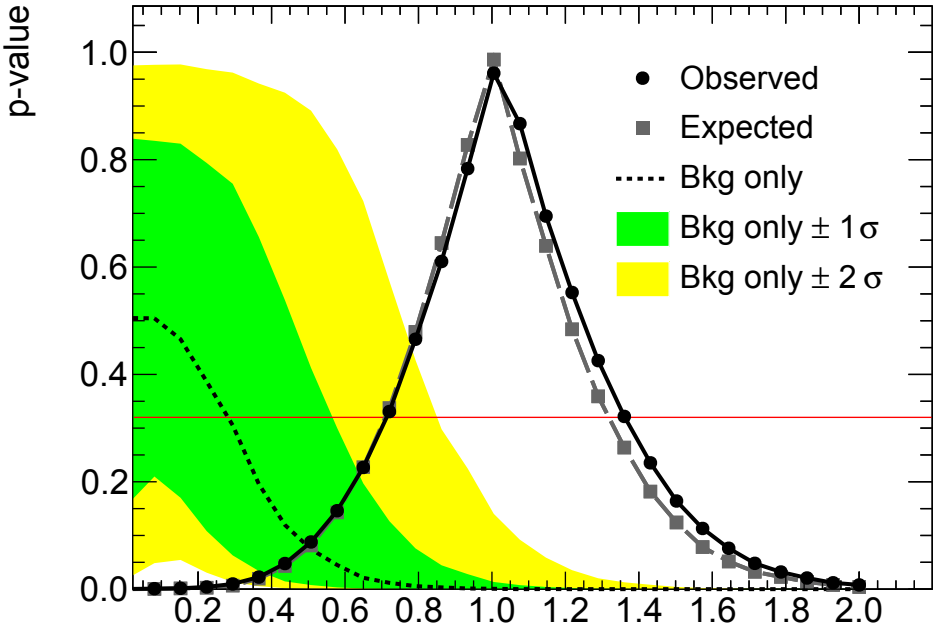


Figure 3.28.: Signal cross section measurement. Hypothesis tests using different test statistics q_μ are performed where the signal hypothesis $\mu = \hat{\mu} = 1.02$ is tested against the background-only hypothesis $\mu = 0$. This results in a curve of p-values from which a 68% coverage central confidence interval is constructed. Again, the observed result is close to the SM expectation. Furthermore, the data can hardly be explained by means of the background-only hypothesis.

way as in case of the discovery test described above. This procedure results in a curve of p-values p_μ. In the neighbourhood of $\mu^{\text{signal}} = \hat{\mu}$ the test statistic q_μ is particularly small and p_μ is large. Outside of this region, q_μ becomes larger and the p-values decrease. A 68% coverage central confidence interval is constructed where the upper and lower limits are given by the signal strengths for which a p-value of $p_\mu = 32\%$ is obtained.

The procedure is performed using the Asimov data as well as the measured data. The results are shown in Fig. 3.28. In case of the Asimov data, one obtains $\mu^{\text{expected}} = 1.00^{+0.32}_{-0.29}$, which corresponds to a total uncertainty of 30%. The central value is given by the MLE $\hat{\mu}$. The observed result is $\mu^{\text{observed}} = 1.02^{+0.34}_{-0.31}$, which corresponds to a total uncertainty of 32%. Once again, the observed result is close to the SM expectation. The precision of the theoretical prediction, which is 4% (Eq. 1.39), is not reached by the measurement. In addition, Fig. 3.28 also shows the expected p-values given the background-only hypothesis. In accordance with the results discussed above, the s-channel signal is needed in order to explain the measured data.

The observed signal strength can be translated into the signal cross section. The MLE and the upper and lower limits simply need to be multiplied with the reference SM cross section provided by Eq. 1.39. The result is

$$\sigma_{s-\text{channel}}^{\text{LHC},\,8\,\text{TeV}} = 5.3^{+1.8}_{-1.6}\,\text{pb}\ . \tag{3.50}$$

In order to understand the impact of different uncertainty components on the total uncertainty, a break-down of the individual contributions is made. More precisely, the determination of the confidence interval is repeated neglecting the sources of uncertainty in question. Consequently, the confidence interval $\Delta\sigma$ becomes smaller. Its difference compared to the total uncertainty is quoted as the impact of the uncertainty in question on the overall uncertainty using the formula

$$\Delta\sigma_{\text{syst}\,i} = \sqrt{(\Delta\sigma_{\text{total}})^2 - (\Delta\sigma_{\text{no syst}\,i})^2}\ . \tag{3.51}$$

This procedure constitutes another point of view on uncertainty contributions compared to the impact of nuisance parameter variations on the MLE of the signal strength discussed above in the context of Fig. 3.23. In addition, it allows one to estimate the impact of uncertainties associated with multiple nuisance parameters like the statistical uncertainty of the predictions or the jet energy scale components. Furthermore, the impact of the finite statistics of the measured data set can be estimated. For this purpose all other uncertainties are neglected and the resulting confidence interval is quoted directly as the uncertainty due to limited data statistics. Instead of pseudo-experiments, approximate formulae are used in order to determine the distributions of the test statistics [199]. The approximation is appropriate as the expected and observed total uncertainties determined in this way, namely 30% and 31%, differ only slightly from the results based on pseudo-experiments, namely 30% and 32%, respectively, as discussed above.

The resulting uncertainty break-down with respect to the Asimov data as well as the measured data is presented in Tab. 3.10. Again the observed results are close to the expectation. The most important single source of uncertainty is given by the amount of data statistics. The most important systematic source of uncertainty is given by the jet energy resolution. The limited number of simulated events is another major contribution. Further important uncertainties are due to the modelling of the t-channel single top-quark background, the renormalization and factorization scale choice regarding the generation

Source	Expected $\Delta\sigma/\sigma$ [%]	Observed $\Delta\sigma/\sigma$ [%]
Data statistics	15	15
Jet energy resolution	10	11
MC statistics	10	11
t-channel modelling	9	10
s-channel generator scales	7	7
b-tagging	7	8
W+jets normalisation	6	7
Jet energy scale	5	5
Luminosity	4	5
t-channel normalisation	4	5
$t\bar{t}$ modelling	2	4
Leptons	3	3
$t\bar{t}$ normalisation	3	3
PDFs	2	2
mistags	1	1
$t\bar{t}$ modelling, generator	1	1
Others	< 1	< 1
Total	30	31

Table 3.10.: Break-down of expected and observed systematic uncertainties. The impact of the data statistics refers to a confidence interval construction where all other uncertainties are neglected. The other uncertainty contributions are estimated according to Eq. 3.51. The precision of these results is on the order of 1%. Asymptotic formulae are used to estimate distributions of the test statistics q_μ. Hence the total uncertainty slightly differs from the result based on pseudo-experiments. The limited data statistics is the most important source of uncertainty. The observed results are close to the expectation.

of s-channel signal events and the b-tagging. These results are in accordance with the ones obtained in the context of the maximum likelihood fit and Fig. 3.23.

Overall, this measurement constitutes a considerable improvement compared to a previous measurement performed by the ATLAS collaboration, where the same data set was used [116, 169]. As mentioned in Sec. 1.3.4, the former measurement used BDTs to separate the s-channel signal (Fig. 1.17d). As a result, a signal significance of 1.3 standard deviations was observed, while the expectation amounted to 1.4 standard deviations.

The uncertainty of the measured cross section was 87%. Thus the precision of the new measurement presented in this thesis surpasses the former result by nearly a factor of three. There are several reasons for this improvement. New calibrations reduced the impact of instrumental uncertainties. Extensions of samples of simulated events led to a decrease of corresponding statistical uncertainties. Furthermore, the replacement of the BDT approach by the TELL method turned out to be another advantage. As the number of simulated W+jets events is relatively small (Tab. A.1 in the appendix), the training of the BDT did not result in an optimal separation of the signal from this background due to an insufficient amount of information. The TELL method, on the other hand, does not involve any training. Thus it provides a better signal separation, in particular with respect to the W+jets background. Given the large uncertainty of the W+jets prediction, this translates into a higher precision of the measurement based on TELLs.

Furthermore, the measurement presented in this thesis can be compared with the corresponding result published by the ATLAS collaboration [190]. The two approaches are identical except for two aspects. The present analysis does not include the W+jets control region in the fit of the model to the data. This is a minor difference. Furthermore, the present analysis refers to s and t-channel single top-quark production cross sections at NLO accuracy, while the result of the ATLAS collaboration refers to cross sections at approximate NNLO accuracy (Sec. 1.3.4). In particular, the t-channel background cross section used by the ATLAS collaboration is 3% larger than the NLO prediction. As a result, the observed signal significance reported by ATLAS amounts to 3.2 instead of 3.4 standard deviations.

Regarding future measurements, the break-down of uncertainties shows how the precision of the measurement could be improved further. Apart from an extension of the amount of data statistics, the most promising approach would be to extend the employed simulation samples. In addition, improvements of the understanding of other major sources of uncertainty are possible. However, given the dominance of the statistical uncertainty and the fact that the data set is fixed as far as pp collisions at a centre-of-mass energy of 8 TeV are concerned, future measurements are most likely to be based on data sets with higher centre-of-mass energies which are yet to be recorded at the LHC. Therefore the potential of the present analysis with regard to the next run of the LHC is estimated in the next section.

3.5. Future prospects

Now that evidence is found for s-channel single top-quark production in pp collisions, future measurements can aim for improvements in terms of sensitivity and precision based on the experiences made in the context of the present analysis. In particular, an increased amount of collision data is desired. Thus the new run of the LHC, which has just started in 2015, is best suited for this purpose. It is supposed to provide large amounts of pp collision data to the ATLAS and CMS collaborations, on the order of 100 inverse femtobarn per experiment. Also its increased centre-of-mass energy of pp collisions, namely 13 TeV, results in an increase of the signal cross section compared to the previous run of the LHC. However, there is also a downside of this new setting. Cross sections of backgrounds including top-quarks increase more strongly with the centre-of-mass energy than the s-channel signal cross section. This is shown in Fig. 3.29. In particular, comparing the centre-of-mass energies of 8 TeV and 13 TeV, the top-quark pair production cross section increases by a factor of 3.3, while the s-channel single top-quark cross section increases only by a factor of 2.0 . In summary, future measurements of s-channel single top-quark production will benefit from increased amounts of data, but the signal event fraction becomes smaller.

In the remaining part of this section, the prospects of s-channel measurements at the LHC are estimated quantitatively. For this purpose, the model employed in the analysis presented above is adjusted. For each process the expected number of events is scaled according to the ratio of its cross sections in pp collisions at the centre-of-mass energies of 13 TeV and 8 TeV. The top-quark pair and single top-quark production cross sections are computed at NNLO and NLO in QCD, respectively, using the HatHor program [70, 108]. The W+jets contribution is scaled according to the $W + b\bar{b}$ cross section computed at NLO in QCD using the MCFM program [187]. The Z+jets and di-boson contribution is scaled according to the $Z + b\bar{b}$ cross section computed at LO using the MCFM program. In case of the prompt lepton fake background the cross section scaling is ignored. However, this choice does not affect the validity of the present study since the impact of this background on the final result is negligible, as shown in Tab. 3.10. Apart from cross sections, also the integrated luminosity of the data set is scaled. Different values are used in order to be able to determine target luminosities necessary to reach measurement goals. Using this scaling of luminosity and cross sections, the expected TELL discriminant distribution is used in order to determine expected results regarding the signal significance and the signal cross section uncertainty. It

is assumed that a sufficient number of simulated events will be available, so that the uncertainty originating from limited simulation statistics can be neglected. All other analysis aspects remain untouched. In summary, the results presented below involve the following approximations:

- For each scattering process the selection efficiencies remain the same compared to the $\sqrt{s} = 8\,\text{TeV}$ analysis presented above.
- Also the impact of each of the systematic uncertainties remains the same.
- The uncertainty due to the limited number of simulated events can be neglected in anticipation of an adequate MC sample production campaign.

Given these assumptions, the following study can only provide a rough estimate of a future analysis performance. However, this is good enough to achieve its aims, which are to assess the future potential of the present analysis approach, to motivate new and improved s-channel single top-quark measurements and to identify interesting target luminosities concerning the new run of the LHC.

As an example, Fig. 3.30 shows the expected distributions of the TELL discriminant at a centre-of-mass energy of 13 TeV in case of an integrated luminosity of $40\,\text{fb}^{-1}$. Clearly the top-quark pair production process becomes even more dominant compared to the $\sqrt{s} = 8\,\text{TeV}$ scenario. On top of the expected distribution, the Asimov data is shown together with Poissonian errors. Their smallness compared to the number of signal events suggests that a significant measurement is possible.

Indeed this is confirmed by a statistical evaluation of the model, which is performed in the same way as in case of the measurement presented in the previous section. The only difference is the fact that asymptotic formulae are used to determine distributions of test statistics. Anyway, the differences to results based on pseudo-experiments are small, as discussed above. The procedure is carried out using different integrated luminosities. Figure 3.31 presents the expected signal significance as a function of the integrated luminosity. A significant identification of the s-channel signal is still possible. The sensitivity of the current measurement can even be surpassed. In particular, an integrated luminosity of about $40\,\text{fb}^{-1}$ is sufficient to allow for the observation of s-channel single top-quark production since the signal significance reaches five standard deviations at this point. Figure 3.32 shows the expected signal cross section uncertainty, which can be reduced considerably compared to the result of the present analysis. At large integrated luminosities, systematic uncertainties become dominant and the total

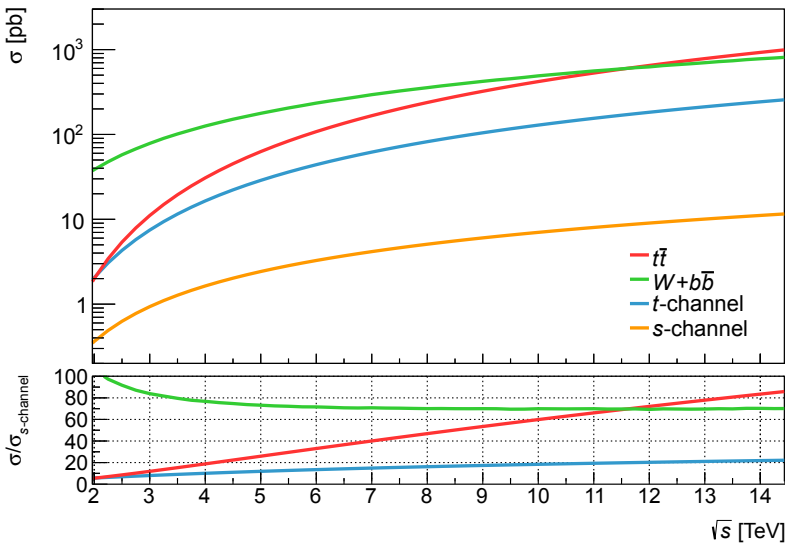


Figure 3.29.: Cross sections relevant to s-channel measurements in pp collisions as a function of the centre-of-mass energy. The top-quark pair production cross section is computed at NNLO in QCD, whereas the s and t-channel single top-quark cross sections are computed at NLO in QCD [70, 108]. The $W+b\bar{b}$ cross section is computed at NLO in QCD as well [187]. The top-quark pair and t-channel single top-quark production cross sections increase more strongly with the centre-of-mass energy than the s-channel single top-quark production cross section.

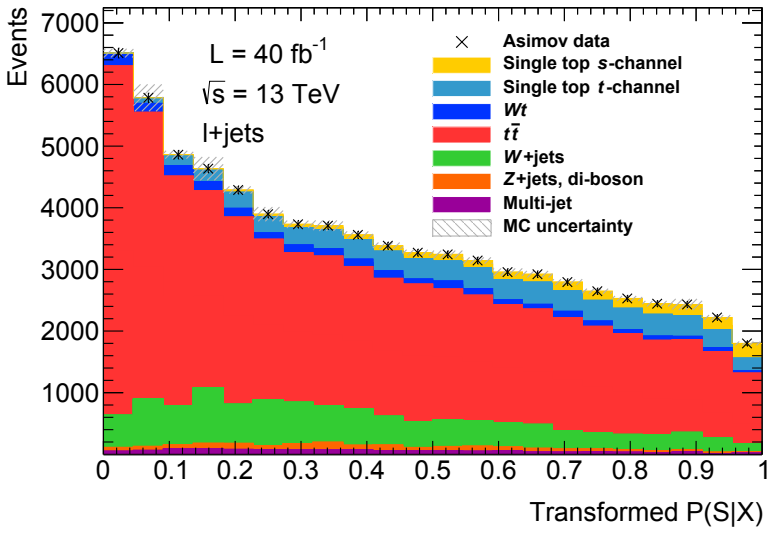


Figure 3.30.: Expected TELL discriminant distribution in case of pp collisions at a centre-of-mass energy of 13 TeV. The integrated luminosity amounts to 40 fb^{-1}. Even though top-quark pair production is by far the dominant process, s-channel single top-quark production is still a significant contribution.

uncertainty is limited to about 15%. Overall, these results motivate further measurements of s-channel single top-quark production using future LHC data sets.

In addition to an increase of the data statistics, further improvements of the analysis are possible. Apart from a better understanding of the modelling of signal and background processes, as well as instrumental uncertainties, there are two aspects which allow one to improve the analysis straightforwardly:

- The number of simulated events should be increased compared to the analysis presented above in order to avoid any significant impact of its statistical uncertainties on the measurement. This requirement should be met not only for the baseline MC event generators but also for those which are used to estimate modelling uncertainties.

- The veto on di-lepton events where the appearance of a second, loosely defined lepton leads to the rejection of the event, should be optimised. It reduces the number of top-quark pair production events significantly, while the number of signal events remains almost the same. The present analysis makes use of the definition of loose leptons which is employed in the context of prompt lepton fake estimations. A detailed study of different loose lepton definitions and their impact on the signal and background event fractions would allow one to optimise the measurement.

In summary, the analysis approach presented above is likely to allow for an observation of s-channel single top-quark production using future LHC data sets, where its full potential can be exploited.

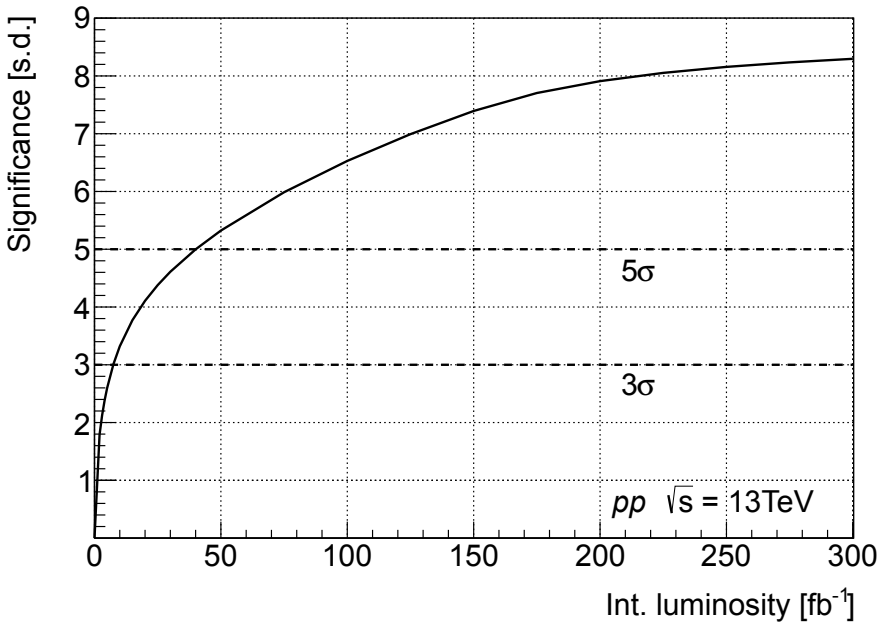


Figure 3.31.: Expected signal significance in pp collisions at a centre-of-mass energy of 13 TeV as a function of the integrated luminosity. Once a data set corresponding to an integrated luminosity of about $40\,\mathrm{fb}^{-1}$ is available, s-channel single top-quark production is expected to be observed with a significance of five standard deviations.

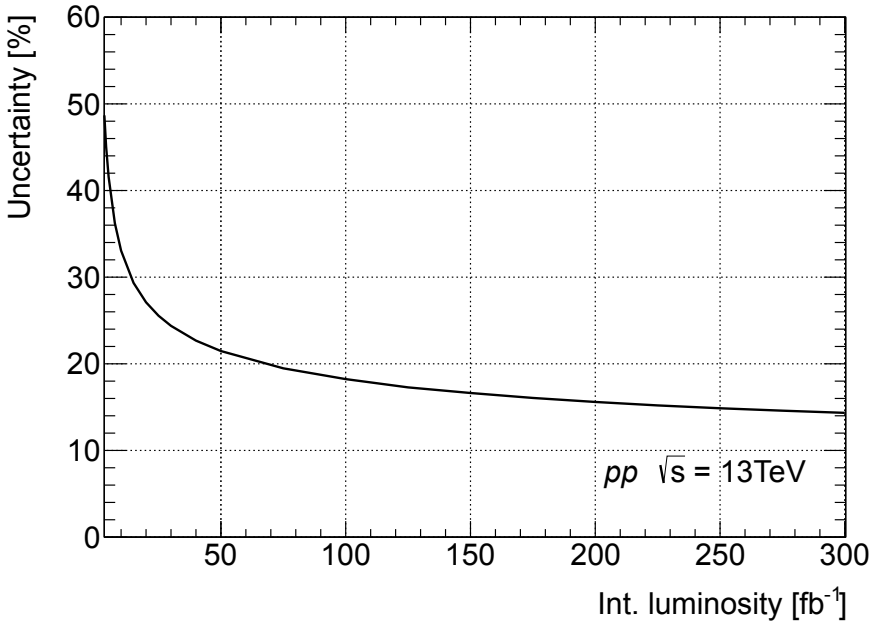


Figure 3.32.: Expected signal cross section uncertainty in pp collisions at a centre-of-mass energy of 13 TeV as a function of the integrated luminosity. Compared to the measurement presented above, the precision can be improved considerably.

Conclusion

In this thesis a search for s-channel single top-quark production at the LHC is performed. This scattering process constitutes an electroweak production of a single top-quark together with a lighter quark. Its interference with the t-channel single top-quark production is small and can be neglected. The SM predicts the occurrence of this process and hence the search for it is a test of the SM. The SM prediction of the s-channel cross section and its theoretical uncertainties are studied in detail. The HatHor program is extended in order to be able to perform the necessary studies efficiently. As a result, the s-channel cross section is predicted with a high precision.

The pp collision data set on which the measurement is based has been recorded with the ATLAS detector at the LHC in 2012 using a centre-of-mass energy of 8 TeV. Due to the complexity of collision events which include top-quarks, all major components of the detector are involved in this measurement. Thanks to a successful operation of the LHC and a high data taking efficiency of ATLAS, a large data set corresponding to an integrated luminosity of $20.3\,\text{fb}^{-1}$ is available. Owing to the rareness of the s-channel single top-quark events, the measurement is restricted to a subset of the data, so that the signal fraction is increased. Selected events contain an isolated electron or muon, two b-tagged jets and missing transverse momentum. All of these objects have high transverse momenta. In order to be able to interpret the measured data, all relevant scattering processes are modelled by means of MC event generators or data-driven estimations. According to the resulting model, the expected number of selected signal events is relatively small, namely 570 out of 15740 events.

Given the smallness of the signal event fraction and the statistical and systematic uncertainties which must be taken into account, sensitivity to the signal process can only be reached if the signal is separated further from its backgrounds. For this purpose the method of total event likelihoods, TELL, is employed. Each collision event is mapped onto a discriminant which approximates the probability of the event to originate from the signal process. Hence signal events result in high values of the TELL discriminant,

whereas background events result in lower values. This separation allows one to identify the presence of possible signal events. The TELL method is implemented in a generic way which makes it applicable to various different measurements. The computations needed to determine discriminants are fast enough to be applied in practice. Contrary to machine learning techniques, no training procedure based on simulated events is involved. In consequence, the separation power provided by the TELL method is not limited by the given number of simulated collision events. Overall, the potential of this method is exploited so that it can be applied easily in order to accomplish a powerful separation of various scattering processes.

Concerning the performed measurement, a statistical model of the distribution of the TELL discriminant is constructed. It is based on the expected signal and background contributions and involves all relevant statistical and systematic uncertainties. Finally, the model is fit to the measured data in order to measure the strength of the s-channel signal. Indeed the signal is very likely to be present since its significance corresponds to 3.4 standard deviations. The measured signal cross section amounts to $\sigma_{s-\text{channel}}^{\text{LHC},\,8\,\text{TeV}} = 5.3^{+1.8}_{-1.6}\,\text{pb}$. These results are in accordance with the SM.

Being the first evidence for s-channel single top-quark production in pp collisions, this measurement improves the understanding of the physics of the top-quark. Furthermore, it motivates new measurements to be performed in the future. An outlook is given with respect to the new run of the LHC where the centre-of-mass energy of pp collisions is raised to 13 TeV. It is expected to observe s-channel single top-quark production with a significance of more than five standard deviations using this new data set. The precision of the measurement is also expected to improve. In summary, the SM is able to explain the measured data, but it remains to be seen if it can continue to do so in the future, not least because a new window to top-quark physics at the LHC is now open.

Appendix A.

Simulation samples

The following tables present the MC event simulation samples used in the measurement. The ATLAS data set identity (DSID) marks the samples. Some properties are provided in addition, including the number of produced events, where event weights are not taken into account. Table A.1 presents the nominal MC samples. The tables A.2 and A.3 present the samples used to study modelling and PDF uncertainties, respectively.

Sample	DSID	Generator	PDF set	N_{MC}
s-channel (ℓ + jets)	110119	Powheg + Pythia 6	CT10	5 995 993
t-channel, top-quark (ℓ + jets)	110090	Powheg + Pythia 6	CT10	4 994 481
t-channel, anti-top-quark (ℓ + jets)	110091	Powheg + Pythia 6	CT10	4 999 879
Wt	110140	Powheg + Pythia 6	CT10	999 692
$t\bar{t}$ (no full-had.)	110404	Powheg + Pythia 6	CT10	49 948 212
$W\to e\nu$, b-quark filter	167740	Sherpa 1.4.1	CT10	14 992 449
$W\to e\nu$, c-quark filter	167741	Sherpa 1.4.1	CT10	5 999 977
$W\to e\nu$, c and b-quark veto	167742	Sherpa 1.4.1	CT10	23 983 938
$W\to \mu\nu$, b-quark filter	167743	Sherpa 1.4.1	CT10	14 990 863
$W\to \mu\nu$, c-quark filter	167744	Sherpa 1.4.1	CT10	5 999 888
$W\to \mu\nu$, c and b-quark veto	167745	Sherpa 1.4.1	CT10	23 997 757
$W\to \tau\nu$, b-quark filter	167746	Sherpa 1.4.1	CT10	14 999 453
$W\to \tau\nu$, c-quark filter	167747	Sherpa 1.4.1	CT10	5 999 680
$W\to \tau\nu$, c and b-quark veto	167748	Sherpa 1.4.1	CT10	23 999 450
$Z\to ee$ + 0 parton	147105	Alpgen + Pythia 6	CTEQ6L1	6 298 988
$Z\to ee$ + 1 partons	147106	Alpgen + Pythia 6	CTEQ6L1	8 184 476
$Z\to ee$ + 2 partons	147107	Alpgen + Pythia 6	CTEQ6L1	3 175 991
$Z\to ee$ + 3 partons	147108	Alpgen + Pythia 6	CTEQ6L1	894 995
$Z\to ee$ + 4 partons	147109	Alpgen + Pythia 6	CTEQ6L1	398 597

Continued on next page

Sample	DSID	Generator	PDF set	N_{MC}
$Z\to ee$ + 5 partons	147110	Alpgen + Pythia 6	CTEQ6L1	229 700
$Z\to\mu\mu$ + 0 parton	147113	Alpgen + Pythia 6	CTEQ6L1	6 298 796
$Z\to\mu\mu$ + 1 partons	147114	Alpgen + Pythia 6	CTEQ6L1	8 193 384
$Z\to\mu\mu$ + 2 partons	147115	Alpgen + Pythia 6	CTEQ6L1	3 175 488
$Z\to\mu\mu$ + 3 partons	147116	Alpgen + Pythia 6	CTEQ6L1	894 799
$Z\to\mu\mu$ + 4 partons	147117	Alpgen + Pythia 6	CTEQ6L1	393 200
$Z\to\mu\mu$ + 5 partons	147118	Alpgen + Pythia 6	CTEQ6L1	229 200
$Z\to\tau\tau$ + 0 parton	147121	Alpgen + Pythia 6	CTEQ6L1	19 202 764
$Z\to\tau\tau$ + 1 partons	147122	Alpgen + Pythia 6	CTEQ6L1	10 674 582
$Z\to\tau\tau$ + 2 partons	147123	Alpgen + Pythia 6	CTEQ6L1	3 765 893
$Z\to\tau\tau$ + 3 partons	147124	Alpgen + Pythia 6	CTEQ6L1	1 096 994
$Z\to\tau\tau$ + 4 partons	147125	Alpgen + Pythia 6	CTEQ6L1	398 798
$Z\to\tau\tau$ + 5 partons	147126	Alpgen + Pythia 6	CTEQ6L1	229 799
$Z\to ee + c\bar{c}$ + 0 parton	200432	Alpgen + Pythia 6	CTEQ6L1	284 999
$Z\to ee + c\bar{c}$ + 1 partons	200433	Alpgen + Pythia 6	CTEQ6L1	499 500
$Z\to ee + c\bar{c}$ + 2 partons	200434	Alpgen + Pythia 6	CTEQ6L1	498 997
$Z\to ee + c\bar{c}$ + 3 partons	200435	Alpgen + Pythia 6	CTEQ6L1	443 697
$Z\to\mu\mu + c\bar{c}$ + 0 parton	200440	Alpgen + Pythia 6	CTEQ6L1	298 998
$Z\to\mu\mu + c\bar{c}$ + 1 partons	200441	Alpgen + Pythia 6	CTEQ6L1	499 799
$Z\to\mu\mu + c\bar{c}$ + 2 partons	200442	Alpgen + Pythia 6	CTEQ6L1	499 500
$Z\to\mu\mu + c\bar{c}$ + 3 partons	200443	Alpgen + Pythia 6	CTEQ6L1	443 999
$Z\to\tau\tau + c\bar{c}$ + 0 parton	200448	Alpgen + Pythia 6	CTEQ6L1	299 000
$Z\to\tau\tau + c\bar{c}$ + 1 partons	200449	Alpgen + Pythia 6	CTEQ6L1	199 998
$Z\to\tau\tau + c\bar{c}$ + 2 partons	200450	Alpgen + Pythia 6	CTEQ6L1	99 800
$Z\to\tau\tau + c\bar{c}$ + 3 partons	200451	Alpgen + Pythia 6	CTEQ6L1	49 400
$Z\to ee$ + bb + 0 parton	200332	Alpgen + Pythia 6	CTEQ6L1	1 799 992
$Z\to ee$ + bb + 1 partons	200333	Alpgen + Pythia 6	CTEQ6L1	999 896
$Z\to ee$ + bb + 2 partons	200334	Alpgen + Pythia 6	CTEQ6L1	994 594
$Z\to ee$ + bb + 3 partons	200335	Alpgen + Pythia 6	CTEQ6L1	885 392
$Z\to\mu\mu$ + bb + 0 parton	200340	Alpgen + Pythia 6	CTEQ6L1	1 799 797
$Z\to\mu\mu$ + bb + 1 partons	200341	Alpgen + Pythia 6	CTEQ6L1	999 897
$Z\to\mu\mu$ + bb + 2 partons	200342	Alpgen + Pythia 6	CTEQ6L1	999 395
$Z\to\mu\mu$ + bb + 3 partons	200343	Alpgen + Pythia 6	CTEQ6L1	880 894
$Z\to\tau\tau$ + bb + 0 parton	200348	Alpgen + Pythia 6	CTEQ6L1	300 000
$Z\to\tau\tau$ + bb + 1 partons	200349	Alpgen + Pythia 6	CTEQ6L1	100 000
$Z\to\tau\tau$ + bb + 2 partons	200350	Alpgen + Pythia 6	CTEQ6L1	50 000
$Z\to\tau\tau$ + bb + 3 partons	200351	Alpgen + Pythia 6	CTEQ6L1	49 800

Continued on next page

Sample	DSID	Generator	PDF set	N_{MC}
WW	105985	Herwig	CTEQ6L1	2 499 890
ZZ	105986	Herwig	CTEQ6L1	245 000
WZ	105987	Herwig	CTEQ6L1	999 998

Table A.1.: All nominal MC samples used in the present analysis. The lepton sign ℓ denotes either e, μ or τ.

Sample	DSID	Generator	PDF set	N_{MC}
s-channel (ℓ + jets)	110120	aMC@NLO + Herwig	CT10	2 964 982
s-channel (ℓ + jets), μ_F =2.0, μ_R =2.0, radHi	110040	Powheg + Pythia 6	CT10	1 000 000
s-channel (ℓ + jets), μ_F =2.0, μ_R =1.0	110041	Powheg + Pythia 6	CT10	999 999
s-channel (ℓ + jets), μ_F =1.0, μ_R =2.0, radHi	110042	Powheg + Pythia 6	CT10	1 000 000
s-channel (ℓ + jets), μ_F =0.5, μ_R =1.0	110043	Powheg + Pythia 6	CT10	999 996
s-channel (ℓ + jets), μ_F =1.0, μ_R =0.5, radLo	110044	Powheg + Pythia 6	CT10	1 000 000
s-channel (ℓ + jets), μ_F =0.5, μ_R =0.5, radLo	110045	Powheg + Pythia 6	CT10	999 999
t-channel (ℓ + jets)	110095	aMC@NLO + Herwig	CT10	999 896
Wt	108346	MC@NLO + Herwig	CT10	1 999 194
$t\bar{t}$ (no full-had.)	105200	MC@NLO + Herwig	CT10	14 997 103
$t\bar{t}$ (no full-had.)	110407	Powheg + Pythia	CT10	14994480
$t\bar{t}$ (no full-had.)	110408	Powheg + Pythia	CT10	14990989
$W\to e\nu$ + 0 parton	147025	Alpgen + Pythia 6	CTEQ6L1	29 464 244
$W\to e\nu$ + 1 partons	147026	Alpgen + Pythia 6	CTEQ6L1	47 936 004
$W\to e\nu$ + 2 partons	147027	Alpgen + Pythia 6	CTEQ6L1	17 495 947
$W\to e\nu$ + 3 partons	147028	Alpgen + Pythia 6	CTEQ6L1	4 855 289
$W\to e\nu$ + 4 partons	147029	Alpgen + Pythia 6	CTEQ6L1	5 403 283
$W\to e\nu$ + 5 partons	147030	Alpgen + Pythia 6	CTEQ6L1	2 787 277
$W\to \mu\nu$ + 0 parton	147033	Alpgen + Pythia 6	CTEQ6L1	31 965 655
$W\to \mu\nu$ + 1 partons	147034	Alpgen + Pythia 6	CTEQ6L1	43 622 615
$W\to \mu\nu$ + 2 partons	147035	Alpgen + Pythia 6	CTEQ6L1	17 611 454

Continued on next page

Sample	DSID	Generator	PDF set	N_{MC}
$W\to\mu\nu$ + 3 partons	147036	Alpgen + Pythia 6	CTEQ6L1	4 796 077
$W\to\mu\nu$ + 4 partons	147037	Alpgen + Pythia 6	CTEQ6L1	5 498 881
$W\to\mu\nu$ + 5 partons	147038	Alpgen + Pythia 6	CTEQ6L1	2 790 985
$W\to\tau\nu$ + 0 parton	147041	Alpgen + Pythia 6	CTEQ6L1	31 877 158
$W\to\tau\nu$ + 1 partons	147042	Alpgen + Pythia 6	CTEQ6L1	48 070 179
$W\to\tau\nu$ + 2 partons	147043	Alpgen + Pythia 6	CTEQ6L1	17 586 943
$W\to\tau\nu$ + 3 partons	147044	Alpgen + Pythia 6	CTEQ6L1	4 982 982
$W\to\tau\nu$ + 4 partons	147045	Alpgen + Pythia 6	CTEQ6L1	2 553 295
$W\to\tau\nu$ + 5 partons	147046	Alpgen + Pythia 6	CTEQ6L1	794 096
$W\to\ell\nu + b\bar{b}$ + 0 parton	200256	Alpgen + Pythia 6	CTEQ6L1	1 599 997
$W\to\ell\nu + b\bar{b}$ + 1 partons	200257	Alpgen + Pythia 6	CTEQ6L1	1 398 396
$W\to\ell\nu + b\bar{b}$ + 2 partons	200258	Alpgen + Pythia 6	CTEQ6L1	699 398
$W\to\ell\nu + b\bar{b}$ + 3 partons	200259	Alpgen + Pythia 6	CTEQ6L1	398 397
$W\to\ell\nu + c\bar{c}$ + 0 parton	200156	Alpgen + Pythia 6	CTEQ6L1	4 299 592
$W\to\ell\nu + c\bar{c}$ + 1 partons	200157	Alpgen + Pythia 6	CTEQ6L1	4 137 891
$W\to\ell\nu + c\bar{c}$ + 2 partons	200158	Alpgen + Pythia 6	CTEQ6L1	2 394 394
$W\to\ell\nu + c\bar{c}$ + 3 partons	200159	Alpgen + Pythia 6	CTEQ6L1	985 295
$W\to\ell\nu + c$ + 0 parton	200056	Alpgen + Pythia 6	CTEQ6L1	22 999 046
$W\to\ell\nu + c$ + 1 partons	200057	Alpgen + Pythia 6	CTEQ6L1	8 198 769
$W\to\ell\nu + c$ + 2 partons	200058	Alpgen + Pythia 6	CTEQ6L1	2 090 290
$W\to\ell\nu + c$ + 3 partons	200059	Alpgen + Pythia 6	CTEQ6L1	499 498
$W\to\ell\nu + c$ + 4 partons	200060	Alpgen + Pythia 6	CTEQ6L1	199 499

Table A.2.: MC samples used in order to estimate modelling uncertainties. The lepton sign ℓ denotes either e, μ or τ.

Sample	DSID	Generator	PDF set	N_{MC}
s-channel (e + jets)	108343	MC@NLO + Herwig	CT10	199 997
s-channel (μ + jets)	108344	MC@NLO + Herwig	CT10	200 000
s-channel (τ + jets)	108345	MC@NLO + Herwig	CT10	199 999
t-channel (ℓ + jets)	110101	AcerMC + Pythia 6	CTEQ6L1	8 997 672
Wt	108346	MC@NLO + Herwig	CT10	1 999 194
$t\bar{t}$ (no full-had.)	105200	MC@NLO + Herwig	CT10	14 997 103

Table A.3.: Top-quark MC samples used for the evaluation of PDF uncertainties. For all other processes, the nominal samples are used. The lepton sign ℓ denotes either e, μ or τ.

Sample	DSID	Generator	N_{MC}
$t\bar{t}$	105200	MC@NLO + Herwig	14983835
$t\bar{t}$	117200	MC@NLO + Herwig	9990989
t-channel (e+jets)	117360	AcerMC + Pythia	1999999
t-channel (μ+jets)	117361	AcerMC + Pythia	1990999

Table A.4.: MC samples used to determine transfer functions. The $t\bar{t}$ samples are used in case of resolution studies provided by the KLFitter package [177]. The t-channel single top-quark samples, on the other hand, are used in case of reconstruction efficiency studies.

Appendix B.

Single top-quark cross section computations

Various theoretical studies of single top-quark production are available, as pointed out in Sec. 1.3.4. Regarding single top-quark measurements, thorough studies related to all of the three production channels are provided by experimental collaborations and improvements of these analyses are ongoing. While the Tevatron has finished operation and the CDF and D0 collaborations presented a final combination of their single top-quark cross section measurements [80], the LHC has just started to explore physics at the TeV scale. The ATLAS and CMS collaborations will analyse more and more data and improve their measurement techniques. Tests of SM predictions of single top-quark production are a vital part of their research program. Furthermore, single top-quark production constitutes a background for rare processes which become accessible at the LHC. Thus, a solid theoretical understanding of single top-quark production is mandatory. In terms of differential cross sections, several tools exist which allow one to perform calculations at NLO accuracy. At the parton level, the MCFM program serves this purpose [62–64, 187], while the Monte Carlo generators aMC@NLO and Powheg allow one to combine NLO calculations with parton showers [95–97, 127].

The aforementioned tools can also be used to calculate total cross sections. However, being designed to compute differential ones, their calculations incorporate phase space integrations of various final state particles. In consequence, these total cross section calculations are relatively expensive in terms of computing time. Detailed studies of total cross sections can involve numerous of these calculations, where renormalization and factorization scales, PDFs and parameters like particle masses are varied. In order to simplify these studies, the HatHor program is extended, so that it allows for the

calculation of single top-quark production cross sections based on the aforementioned NLO calculations. This includes all of the three channels of single top-quark production. Being developed for the case of top-quark pair production originally [108], the extended HatHor program also makes use of total single top-quark production cross sections at the parton level. Thus the time consuming phase space integration is avoided and only integrations over the initial parton momentum fractions x_1 and x_2 remain.

The extension of the HatHor program to single top-quark production is validated extensively. For numerous choices of parameters, hadronic cross sections computed with HatHor are compared with corresponding results of MCFM. In all cases, they agree with each other up to negligible deviations on the order of 10^{-3}. Now HatHor can be used to obtain total single top-quark production cross sections and to study their dependence on various parameters. For instance, it can be used in the context of fits to single top-quark measurements like PDF fits which include t-channel single top-quark cross sections in order to constrain the b-quark PDF of the proton. Furthermore, the program provides a public access to the most recent complete fixed order calculations of total single top-quark production cross sections, which allows experimentalists to compare their results with SM predictions in a transparent way. While the capabilities of the HatHor program are presented above in Sec. 1.3.4, the following section focuses on the way in which the HatHor program works.

B.1. Partonic cross sections at NLO in QCD

As explained in Sec. 1.2.1, hadronic cross sections are given by convolutions of partonic cross sections with PDFs. In HatHor, the integration over the initial parton momentum fractions x_1 and x_2 is based on the Vegas algorithm [66, 67], while PDF sets are obtained via the Les Houches Accord PDF Interface (LHAPDF) [186]. The main task concerning the creation of the HatHor program is the determination of the partonic cross sections $\hat{\sigma}_{ij}(\hat{s},\, \mu_R,\, \mu_F)$. For each production channel, they can be expanded in the strong coupling α_S, which depends on the renormalization scale μ_R. The expansion up to NLO reads

$$\hat{\sigma}_{ij}(\hat{s},\, \mu_R,\, \mu_F) = \alpha_S^k(\mu_R)\hat{\sigma}_{ij}^{(0)}(\hat{s}) + \alpha_S^{k+1}(\mu_R)\hat{\sigma}_{ij}^{(1)}(\hat{s},\, \mu_R,\, \mu_F) + \mathcal{O}(\alpha_S^{k+2}) \,. \tag{B.1}$$

In case of the s and t-channel processes $k = 0$, while in case of the associated Wt production $k = 1$ (Fig. 1.11). The leading order cross sections $\hat{\sigma}_{ij}^{(0)}$ can be derived analytically. Beyond the leading order, cross section computations are much more involved and usually there are no compact analytic expressions. Corrections of virtual

and real radiation as well as contributions of the factorisation of initial state singularities need to be taken into account. For instance, the NLO correction schematically reads

$$\begin{aligned}\alpha_S^{k+1}(\mu_R)\hat{\sigma}_{ij}^{(1)} &= \hat{\sigma}_{ij}^{V} + \hat{\sigma}_{ij}^{R} + \hat{\sigma}_{ij}^{\mathrm{fac}} \\ &= \int \mathrm{d}\Phi_n \frac{\mathrm{d}\hat{\sigma}_{ij}^{V}}{\mathrm{d}\Phi_n} + \int \mathrm{d}\Phi_{n+1} \frac{\mathrm{d}\hat{\sigma}_{ij}^{R}}{\mathrm{d}\Phi_{n+1}} + \int \mathrm{d}x \mathrm{d}\Phi_n \frac{\mathrm{d}\hat{\sigma}_{ij}^{\mathrm{fac},\,x}}{\mathrm{d}\Phi_n} \,, \end{aligned} \tag{B.2}$$

where $\mathrm{d}\Phi_n$ denotes the n particle phase space measure and the integration over the variable x indicates a convolution introduced by the factorisation of initial state singularities. Each of the contributions contains singularities, but they cancel in the sum. One way to achieve this cancellation is to employ the Catani-Seymour subtraction formalism [200, 201], where appropriate terms are added and subtracted in order to render each contribution finite while the sum is not changed. In practice these calculations involve numerical integrations.

A complete computer code which allows one to perform these NLO calculations regarding all of the three single top-quark production channels is published in the MCFM program [187]. It is designed for the computation of hadronic cross sections. Clearly, the partonic cross sections can be extracted from the program code. However, due to the complexity of the NLO calculations modifications are potentially error-prone. For this reason, three complementary approaches are chosen in order to determine NLO corrections. Their results are checked against each other in order to validate them.

One approach is to avoid any intervention concerning the code of the NLO corrections and to use the interface to PDFs instead. Pseudo-PDFs allow for an extraction of the partonic cross sections. For the partonic channel of interest, narrow Gaussian distributions are used as parton distribution functions to probe the partonic cross section $\hat{\sigma}_{ij}$ at the partonic centre-of-mass energy $\sqrt{\hat{s}} = x_0\sqrt{s}$. More precisely,

$$f_{i,j}(x_0,\, \mu_0) = \frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{(x-x_0)^2}{2\delta^2}\right) \,, \tag{B.3}$$

where δ denotes the width of the distribution. In the course of the determination of $\hat{\sigma}_{ij}$ only the PDFs f_i, f_j are chosen in this way, while all other PDFs vanish. The renormalization and factorization scales are fixed at $\mu_R = \mu_F = \mu_0$. The scale dependencies of the higher order corrections are restored by means of the renormalization group equation, as

discussed below. Using the pseudo-PDFs, one obtains to a good approximation

$$\sigma_{\text{had}}(s,\, \mu_0,\, \mu_0)|_{\text{pseudo-PDFs}} = \hat{\sigma}_{ij}(x_0^2 s,\, \mu_0,\, \mu_0) + \mathcal{O}(\delta^2)\ . \tag{B.4}$$

The uncertainty introduced by this method is proportional to the square of the width of the pseudo-PDFs. Another uncertainty arises at the kinematic threshold where a part of the pseudo-PDF is essentially cut off, so that it is not normalised to unity. As these limitations are small, a precise extraction of $\hat{\sigma}_{ij}$ is possible, given a choice of a small width δ. This approach is validated successfully by means of reproductions of well known parton level results, namely top-quark pair production cross sections and analytic cross sections of single top-quark production at the leading order. A drawback of this method is the computational effort needed to perform the integration. Apart from that it is universally applicable and avoids the need to modify the source program.

Alternatively, the MCFM program is adjusted in order to remove the integration over PDFs [104]. The remaining phase space and convolution integrals lead to the desired partonic cross sections. This approach is numerically more efficient than the usage of pseudo-PDFs. However, the necessity for considerable changes of the original program comes as a disadvantage. The results of this approach are compared extensively with those of the pseudo-PDF approach. Deviations are found to be less than 10^{-3}, which proves the correctness of both methods. As a third approach, partonic cross sections of s and t-channel production at NLO are implemented according to Ref. [65] [202]. The results agree with those of the other approaches.

The cross sections depend on various parameters which are not fixed by the SM but are constrained more or less tightly by measurements. In HatHor, the current world average values of the relevant parameters are chosen by default. They can be changed as desired, except for the W boson mass, which is tightly constrained by measurements. In particular, different top-quark masses according to the pole mass scheme can be chosen. For this purpose, the partonic cross sections are determined on a two-dimensional grid of several top-quark masses and partonic centre-of-mass energies. Within the computations of hadronic cross sections, partonic cross sections are determined by means of interpolations on this grid. Details on this procedure can be found in the dedicated publication [70]. As a result, the top-quark mass can be adjusted within the range $165\,\text{GeV} < m_t < 950\,\text{GeV}$, which allows one not only to study mass dependencies of cross sections within the SM, but also to compute electroweak production cross sections of hypothetical heavier versions of the top-quark.

As HatHor provides total cross sections, no kinematic cuts are involved. However, there is one exception in the case of associated Wt production. At NLO this process interferes with top-quark pair production as mentioned in Sec. 1.3.2. The impact of the interference can be reduced significantly by a cut on the transverse momentum $p_{\mathrm{T}}^{\bar{b}}$ of the final state anti-b-quark, which is usually large in top-quark pair but small in Wt production. Similar relations hold in case of the charge conjugated process. The MCFM program allows the user to set such a cut. In HatHor, the suggestion given in Ref. [61], $p_{\mathrm{T}}^{\bar{b}} < 25\,\mathrm{GeV}$, is followed.

B.2. Renormalization and factorization scale dependence

Partonic cross sections are determined given one choice of scales as described above. The renormalization group equation allows one to derive the scale dependencies of the higher order corrections. In the following this procedure is explained, focussing on the NLO corrections, while higher orders are treated accordingly [203]. Before the procedure itself is discussed, it is helpful to rearrange the hadronic cross section σ_{had} of Eq. 1.34 by means of the following convolution,

$$\begin{aligned}(f \otimes g)(x) &= \int_0^1\int \mathrm{d}x_1\mathrm{d}x_2\,\delta(x - x_1x_2) f(x_1) g(x_2) \\ &= \int_x^1 \frac{\mathrm{d}y}{y} f\left(\frac{x}{y}\right) g(x)\ . \end{aligned} \tag{B.5}$$

Associativity and commutativity hold for this mapping. Furthermore, it is useful to introduce a scaling parameter ϱ and cross section scaling functions f_{ij},

$$\varrho = \frac{\hat{s}_{\mathrm{min}}}{\hat{s}} = \frac{M^2}{\hat{s}}, \quad \hat{\sigma}_{ij}(\varrho) = \varrho f_{ij}(\varrho)\ , \tag{B.6}$$

where M denotes the kinematic threshold of the process in question, for instance $M = m_t$ in case of s and t-channel single top-quark production. In order to reduce the number of independent cross section evaluations, CKM matrix elements can be factored out of the scaling functions. However, this is not shown explicitly here. With the aid of these definitions, Eq. 1.34 can be converted into a compact expression using the substitution

$(x_1, x_2) \mapsto (x_1, x_1 x_2)$ and $\varrho = M^2/(x_1 x_2 s)$,

$$\sigma_{\text{had}} = \frac{M^2}{s} f_i \otimes f_j \otimes f_{ij} \left(\frac{M^2}{s}\right) \ . \tag{B.7}$$

Here and in the following repeated indices are summed, while combinations of the same type of parton appear only once. In terms of scaling functions, the expansion B.1 reads

$$f_{ij}(\varrho, \mu_F, \mu_R) = \alpha_S^k f_{ij}^{(0)}(\varrho) + \alpha_S^{k+1} f_{ij}^{(1)}(\varrho, \mu_R, \mu_F) + \mathcal{O}(\alpha_S^{k+2}) \ . \tag{B.8}$$

The leading order scaling functions do not depend on μ_R and μ_F. The NLO corrections are known at one set of scales, given the discussion of the previous section,

$$f_{ij}^{(1)}(\varrho, m_t, m_t) = f_{ij}^{(10)}(\varrho) \ . \tag{B.9}$$

Now two requirements enter. Firstly, partonic cross sections are independent of the renormalization scale, if all orders of α_S could be taken into account. Using the beta function of QCD (Eq. 1.32), this independence can be translated into relations between higher order and lower order scaling functions by means of the identity theorem of power series,

$$\begin{aligned} \frac{\mathrm{d} f_{ij}}{\mathrm{d} \ln \mu_R^2} = 0 \ \forall \ \alpha_S \quad \& \quad \frac{\mathrm{d}\alpha_S}{\mathrm{d} \ln \mu_R^2} &= - \underbrace{\frac{1}{4\pi}\left(11 - \frac{2}{3} n_f\right)}_{=\beta_0} \alpha_S^2 + \mathcal{O}(\alpha_S^3) \\ \Rightarrow \quad \frac{\mathrm{d} f_{ij}^{(1)}}{\mathrm{d} \ln \mu_R^2} &= k \beta_0 f^{(0)} \ , \end{aligned} \tag{B.10}$$

where n_f is the number of approximately massless quark flavours. In HatHor $n_f = 5$ is chosen. The second requirement is the independence of the hadronic cross section of the factorization scale, which again is true if all orders of α_S could be taken into account. Using the DGLAP equations 1.33 for the evolution of the PDFs f_i, another relation among higher and lower order scaling functions is obtained,

$$\begin{aligned} \frac{\mathrm{d} f_i}{\mathrm{d} \ln \mu_F^2} = \alpha_S P_{ij} \otimes f_j \quad \& \quad \frac{\mathrm{d}\sigma_{\text{had}}}{\mathrm{d} \ln \mu_F^2} &= 0 \ \forall \ \alpha_S \ \forall \text{ PDF sets} \\ \Rightarrow \quad \frac{\mathrm{d} f_{ij}^{(1)}}{\mathrm{d} \ln \mu_F^2} &= -P_{ki}^{(0)} \otimes f_{kj}^0 - P_{kj}^{(0)} \otimes f_{ik}^0 \ , \end{aligned} \tag{B.11}$$

where $P_{ij}^{(0)}$ denote the QCD splitting functions at the leading order. The NLO corrections $f_{ij}^{(1)}(\varrho, \mu_F, m_R)$ can now be determined given the two differential equations B.10 and B.11 and the boundary condition B.9. The solution is

$$f_{ij}^{(1)}(\varrho, \mu_F, m_R) = f_{ij}^{(10)}(\varrho) + \ln\frac{\mu_F^2}{m_t^2} f^{11}(\varrho) + k\beta_0 \ln\frac{\mu_R^2}{\mu_F^2} f^{(0)}(\varrho) \ , \tag{B.12}$$

where $f^{(11)}$ is given by

$$f_{ij}^{11}(\varrho) = k\beta_0 f_{ij}^{(0)}(\varrho) - P_{li}^0 \otimes f_{lj}^{(0)} - P_{lj}^0 \otimes f_{il}^{(0)} \ . \tag{B.13}$$

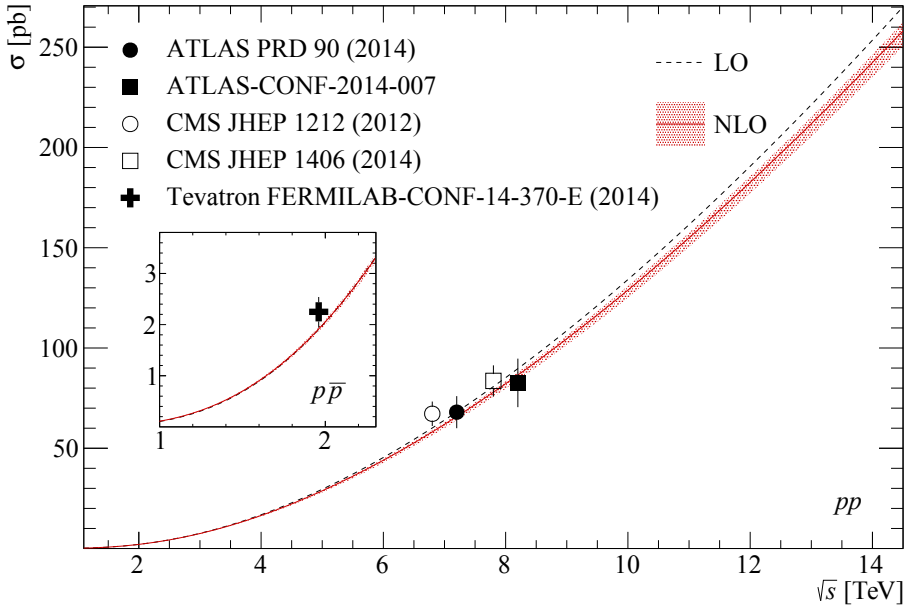
This approach is universally applicable. To give an example, $f_{u\bar{d}}^{(11)}$ in case of s-channel single top-quark production is derived as follows. The process is purely electroweak at the leading order, hence $k = 0$. Regarding the sum over parton types implied by Eq. B.13, gluons do not contribute as the leading order scaling functions $f_{ij}^{(0)}$ vanish if $i = g$ or $j = g$. Furthermore, at the leading order splittings of quarks are given by gluon radiation where the quark flavour propagates on, $P_{q_i q_j}^{(0)} = \delta_{ij} P_{qq}^{(0)}$. These splittings are symmetric with respect to the quark charge, $P_{qq}^{(0)} = P_{\bar{q}\bar{q}}^{(0)}$. Hence the result is

$$f_{u\bar{d}}^{(11)} = -2P_{qq}^{(0)} \otimes f_{u\bar{d}}^{(0)} \ . \tag{B.14}$$

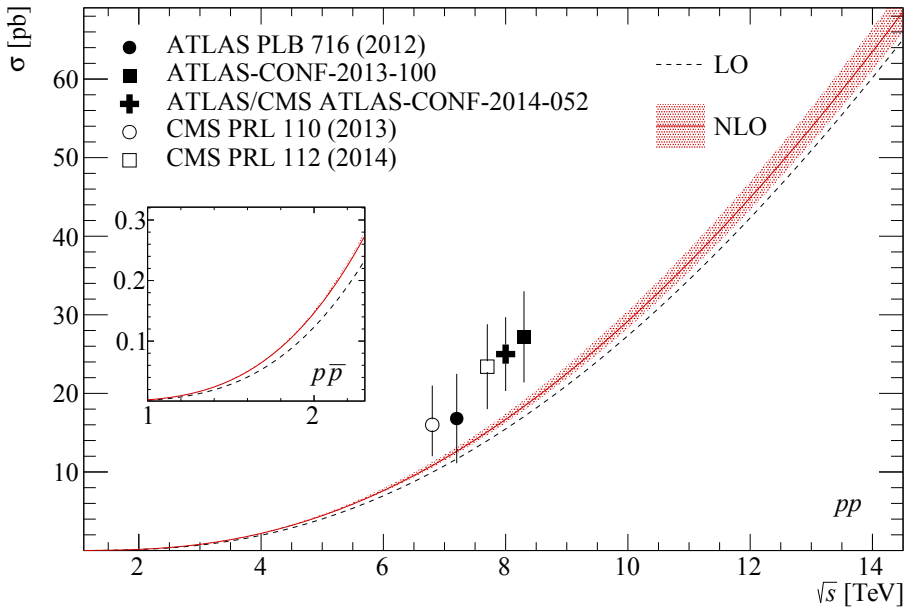
Further results up to NNLO accuracy are presented in the dedicated publication [70].

B.3. Hadronic cross section results

While studies of s-channel single top-quark production cross sections performed with the HatHor program are presented in Sec. 1.3.4, additional results on t-channel and Wt production cross sections are presented in the following. Figure B.1 shows the corresponding cross sections as a function of the centre-of-mass energy for the LHC and the Tevatron, where variations of the scales μ_R and μ_F are taken into account. In both of the production channels, the NLO corrections and scale dependencies are small. In addition, experimental results are shown except for Wt production at the Tevatron, which comes with a particularly small rate. A closer look at the scale dependencies is taken in Fig. B.2, which presents cross section contours obtained from independent variations of the two scales. The residual scale dependencies are covered to a reasonable degree by combined variations of the two scales $\mu = \mu_R = \mu_F$.

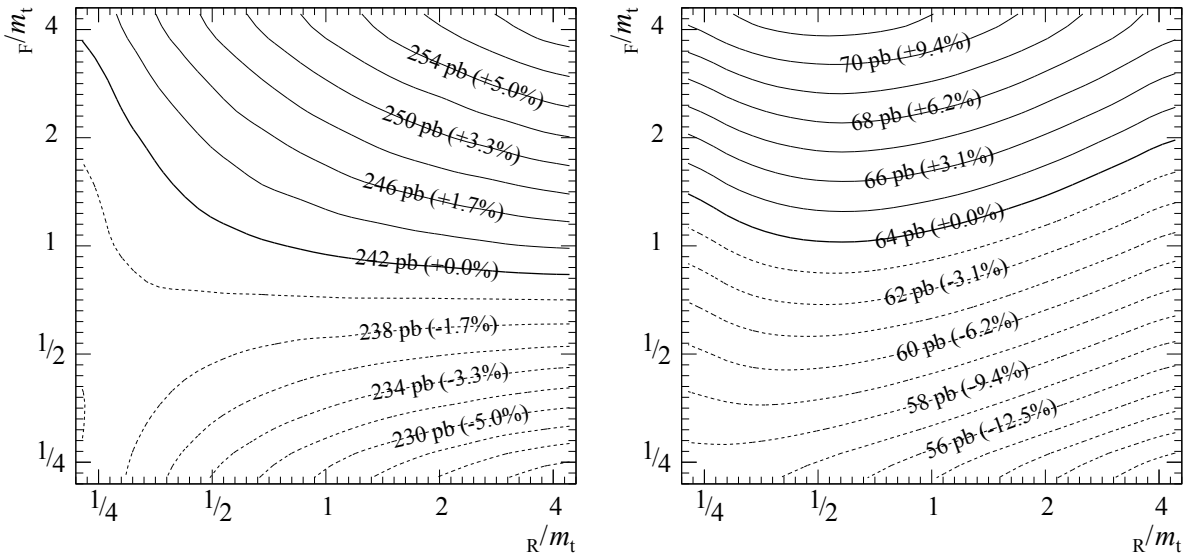


(a) t-channel



(b) Wt production

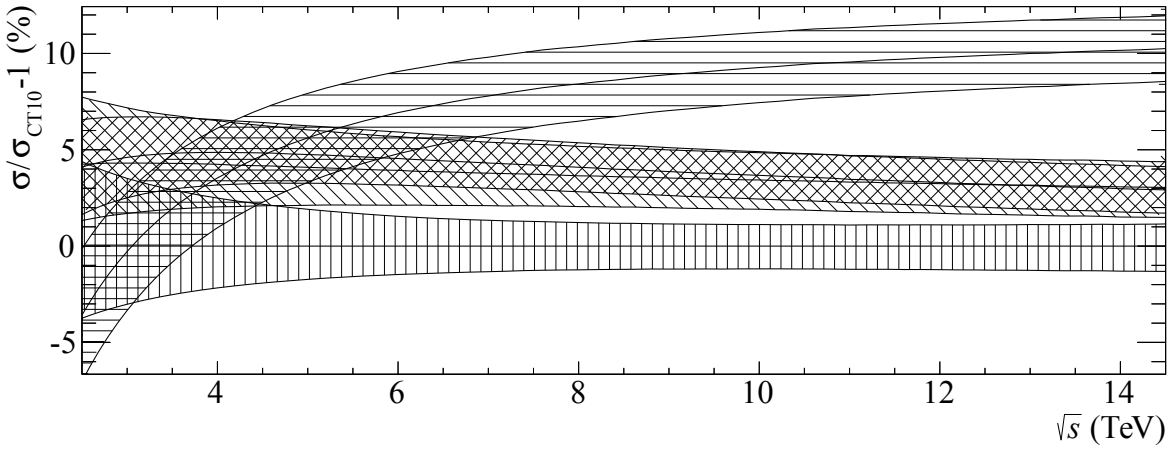
Figure B.1.: Single top-quark production cross sections of t-channel and associated Wt production as a function of the centre-of-mass energy $\sqrt{s}$ in case of pp and $p\bar{p}$ collisions. The calculations are performed at NLO in QCD using the CT10nlo PDF set [56]. The uncertainty bands indicate combined variations of the scales μ_R and μ_F by factors of 2 and 1/2. In addition, measurements from the Tevatron and the LHC are shown [80–82, 204–210].



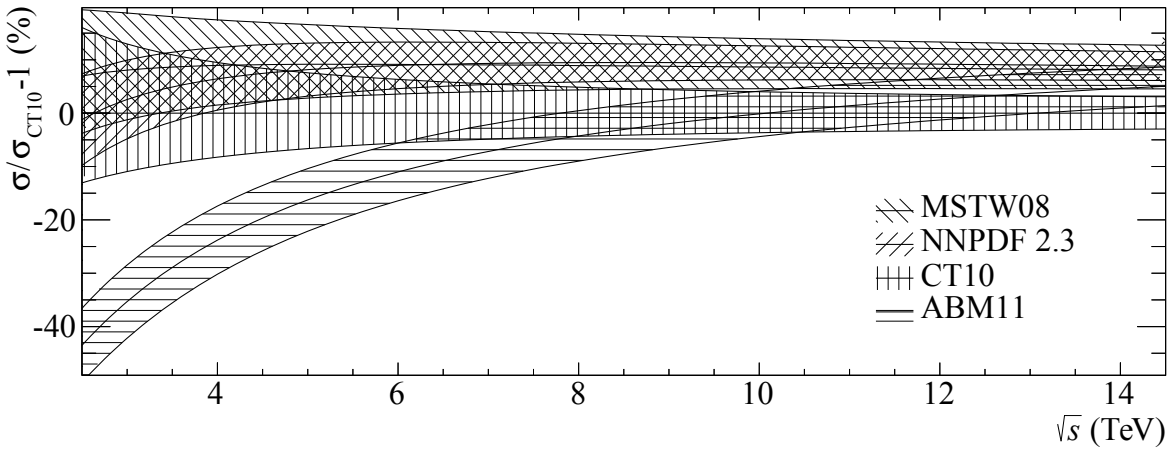
(a) t-channel **(b)** Wt production

Figure B.2.: Scale dependencies of t-channel and Wt single top-quark production cross sections in pp collisions at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$. The calculations are performed at NLO in QCD using the CT10nlo PDF set [56]. In each case the bold contour line indicates the nominal cross section value at $\mu_R = \mu_F = m_t$ [107].

Figure B.3 presents PDF uncertainties of the cross sections. The same PDF sets as in case of the s-channel study, Fig. 1.16, are used. In case of the t-channel, the uncertainties of the individual PDF sets vary between approximately 2.5% at low energies and 1% at high energies. Larger uncertainties are obtained in case of the associated Wt production, which is expected because in this case the gluon PDF, which is hard to constrain, plays an important role. Comparing different PDF sets, CT10 predictions are slightly lower than those obtained using MSTW08 and NNPDF23. Again, as in case of the s-channel, large deviations occur with respect to the ABM11 PDF set.



(a) t-channel



(b) Wt production

Figure B.3.: PDF uncertainties of t-channel and Wt single top-quark production cross sections in pp collisions as a function of the centre-of-mass energy $\sqrt{s}$. PDFs are employed in the same way as in case of the s-channel cross section PDF study presented in Fig. 1.16. Again, the predictions are normalized with respect to the results of the CT10 PDF set. As already observed in the case of the s-channel, differences between PDF sets are larger than the uncertainties of individual sets [107].

Appendix C.

Additional studies

C.1. Standard model fit of control distributions

Fit results

In Sec. 3.1.4 the modelling of the data by means of simulations and data-driven estimates is evaluated in the signal and validation regions by means of dedicated fits of a model of the m_{T}^{W} distribution to the data. Only a reduced set of uncertainties is taken into account, namely the statistical uncertainties of the predictions, theoretical normalisation uncertainties of simulated processes, the normalisation uncertainty of the prompt lepton fake estimate and the uncertainty of the integrated luminosity. This simplified model is sufficient, given the purpose of modelling validation. More precisely, the aim is to investigate the modelling of the background contributions. The determination of the s-channel signal event yield is subject to the final measurement. Therefore, the signal contribution is constrained by its SM prediction in the context of the control distribution fits.

In the following, further information on these modelling studies is provided. The tables C.1, C.2 and C.3 provide the process scale factors η resulting from the control distribution fits according to Eq. 3.42. The electron and muon event selections are combined. Deviations of the fit results from the nominal values $\eta_0 = 1$ are smaller than or equal to the input uncertainties $\Delta\eta_0$ given by the precision of the theoretical cross section predictions. Overall, the resulting scale factors are reasonable, given their limited deviation from the expectation.

Process	η			$(\eta - \eta_0)/\Delta\eta_0$	$\Delta\eta/\Delta\eta_0$
s-channel	1.01	$\pm$	0.05	0.1	0.97
t-channel	1.01	$\pm$	0.04	0.2	0.98
Wt	1.00	$\pm$	0.07	-0.04	1.2
$t\bar{t}$	0.94	$\pm$	0.03	-1.0	0.57
W+jets	1.6	$\pm$	0.2	0.9	0.26
Z + jets & di-boson	1.4	$\pm$	0.9	0.7	1.5
Multi-jet	0.5	$\pm$	0.2	-1.0	0.36

Table C.1.: Control distribution fit results in the signal region. Electron and muon contributions are merged. The scale factors $\eta \pm \Delta\eta$ are determined by a fit to the m_{T}^{W} distribution, where the nominal values $\eta_0 = 1$ and $\Delta\eta_0$ enter corresponding Gaussian constraint terms of the likelihood function. This includes the s-channel signal, as the fit only serves the purpose of validating the modelling of background contributions.

Process	η			$(\eta - \eta_0)/\Delta\eta_0$	$\Delta\eta/\Delta\eta_0$
s-channel	1.00	$\pm$	0.05	0.01	0.97
t-channel	1.00	$\pm$	0.04	0.06	0.97
Wt	1.00	$\pm$	0.07	0.03	1.1
$t\bar{t}$	0.98	$\pm$	0.05	-0.3	0.76
W+light jets	0.84	$\pm$	0.20	-0.5	0.67
W+heavy flavour	1.4	$\pm$	0.1	0.7	0.25
Z + jets & di-boson	0.9	$\pm$	0.8	-0.1	1.3
Multi-jet	0.5	$\pm$	0.2	-1.0	0.34

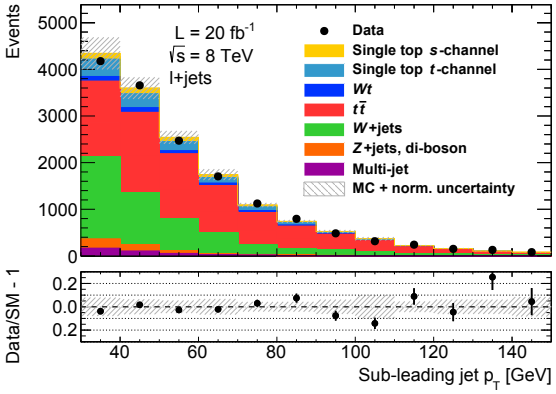
Table C.2.: Control distribution fit results in the W+jets region. Electron and muon contributions are merged. The scale factors $\eta \pm \Delta\eta$ are determined by a fit to the m_{T}^{W} distribution, where the nominal values $\eta_0 = 1$ and $\Delta\eta_0$ enter corresponding Gaussian constraint terms of the likelihood function. This includes the s-channel signal, as the fit only serves the purpose of validating the modelling of background contributions.

Process	η	$(\eta - \eta_0)/\Delta\eta_0$	$\Delta\eta/\Delta\eta_0$
s-channel	1.00 ± 0.05	-0.003	1.0
t-channel	1.00 ± 0.04	-0.005	1.0
Wt	1.00 ± 0.07	-0.01	1.2
$t\bar{t}$	1.03 ± 0.03	0.4	0.55
W+jets	1.2 ± 0.5	0.3	0.76
Z + jets & di-boson	1.0 ± 0.5	0.05	0.76
Multi-jet	0.5 ± 0.2	-1.0	0.50

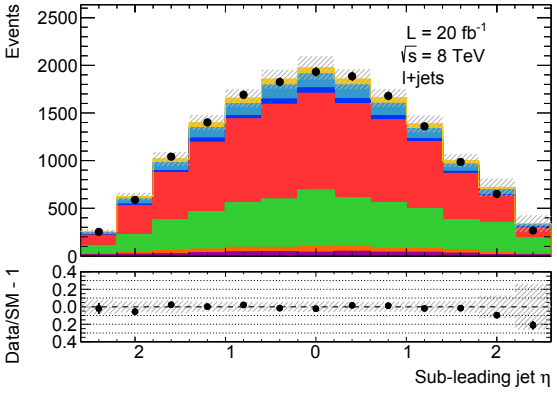
Table C.3.: Control distribution fit results in the $t\bar{t}$ region. Electron and muon contributions are merged. The scale factors $\eta \pm \Delta\eta$ are determined by a fit to the m_{T}^{W} distribution, where the nominal values $\eta_0 = 1$ and $\Delta\eta_0$ enter corresponding Gaussian constraint terms of the likelihood function. This includes the s-channel signal, as the fit only serves the purpose of validating the modelling of background contributions.

Additional control distributions

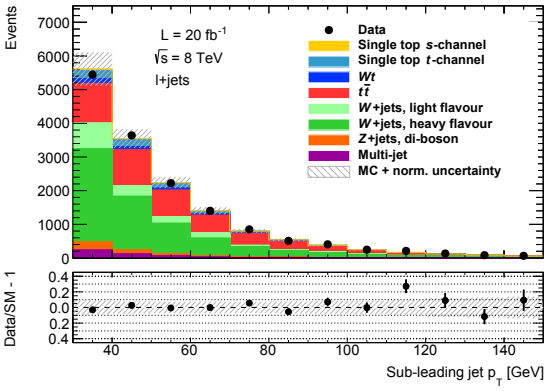
Figure C.1 presents control distributions of the sub-leading jet p_{T} and η in all of the three kinematic regions. The figures C.2 and C.3 present p_{T} and η control distributions of electrons and muons, respectively. The normalisation of each process is scaled according to the corresponding control distribution fit result. In all cases a good modelling of the data is achieved.



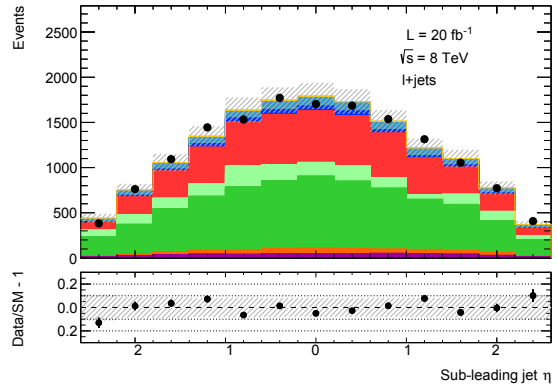
(a) Signal region



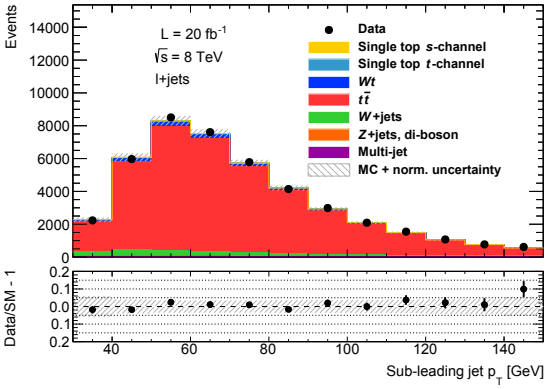
(b) Signal region



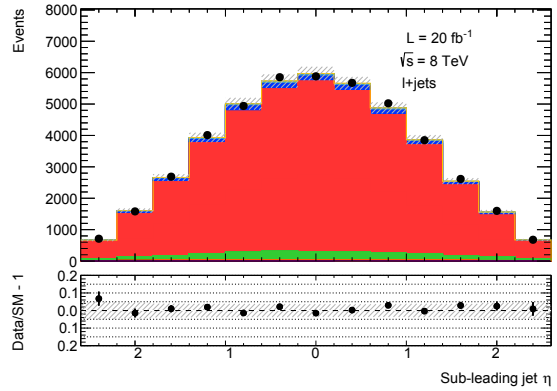
(c) W+jets region



(d) W+jets region

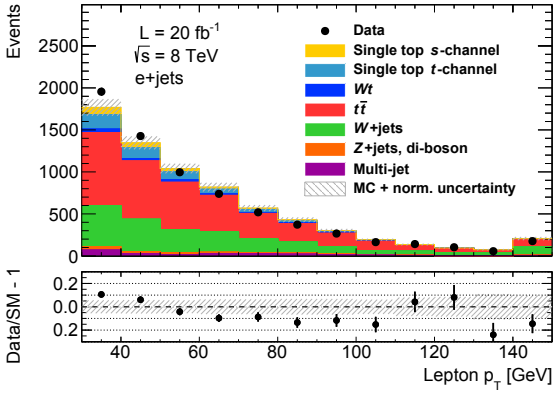


(e) $t\bar{t}$ region

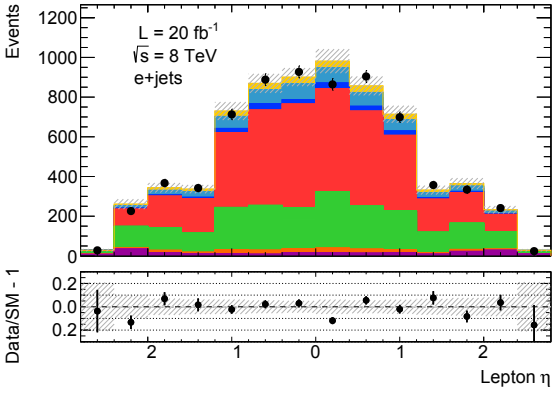


(f) $t\bar{t}$ region

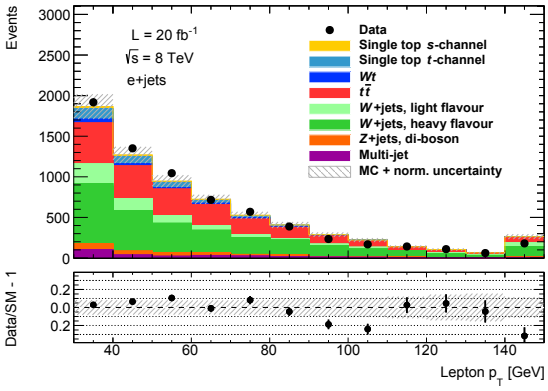
Figure C.1.: Sub-leading jet control distributions. Each process is scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands include the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.



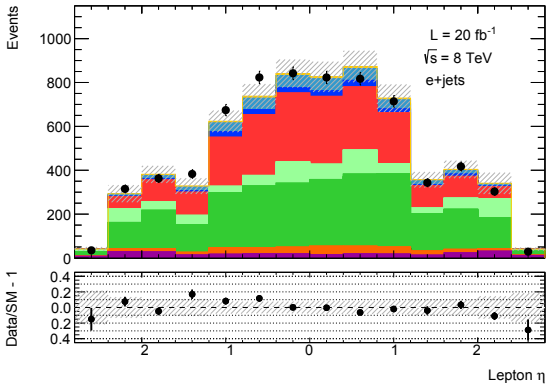
(a) Signal region



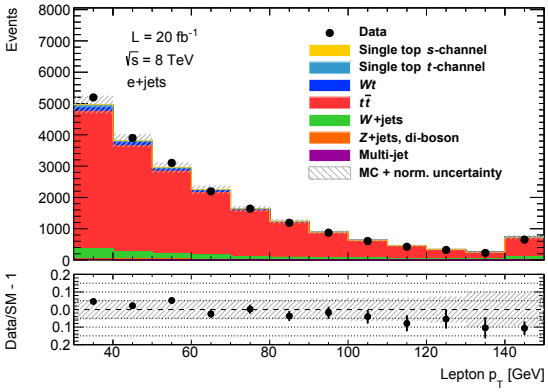
(b) Signal region



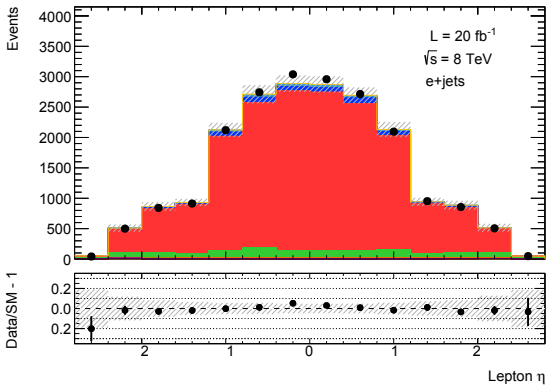
(c) W+jets region



(d) W+jets region

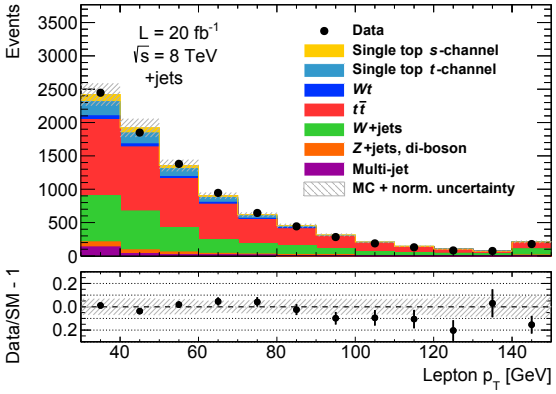


(e) $t\bar{t}$ region

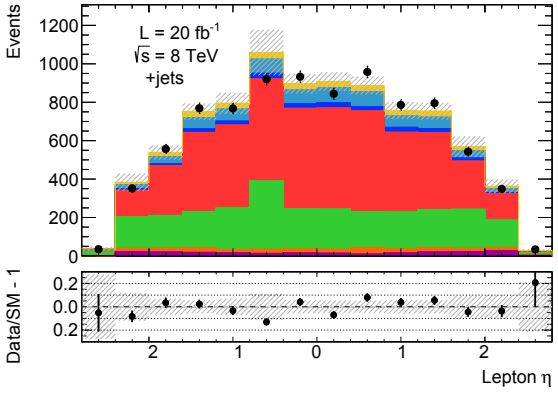


(f) $t\bar{t}$ region

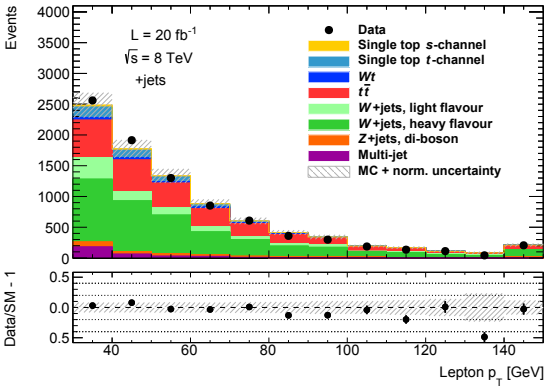
Figure C.2.: Electron control distributions. Each process is scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands include the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.



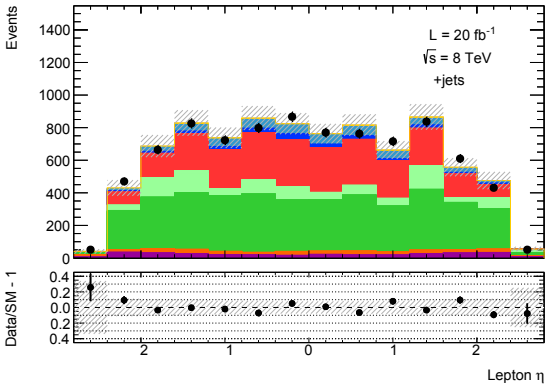
(a) Signal region



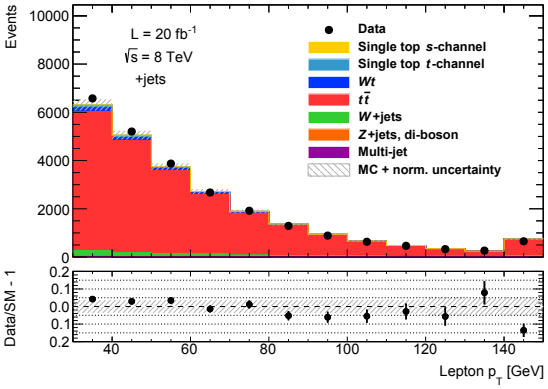
(b) Signal region



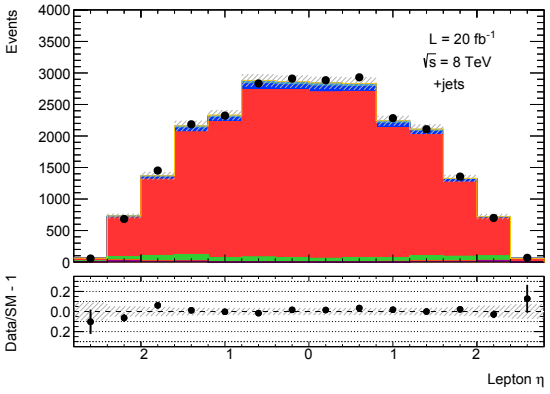
(c) W+jets region



(d) W+jets region



(e) $t\bar{t}$ region



(f) $t\bar{t}$ region

Figure C.3.: Muon control distributions. Each process is scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. The uncertainty bands include the statistical uncertainties of each process as well as their normalisation uncertainties which are constrained by the fit.

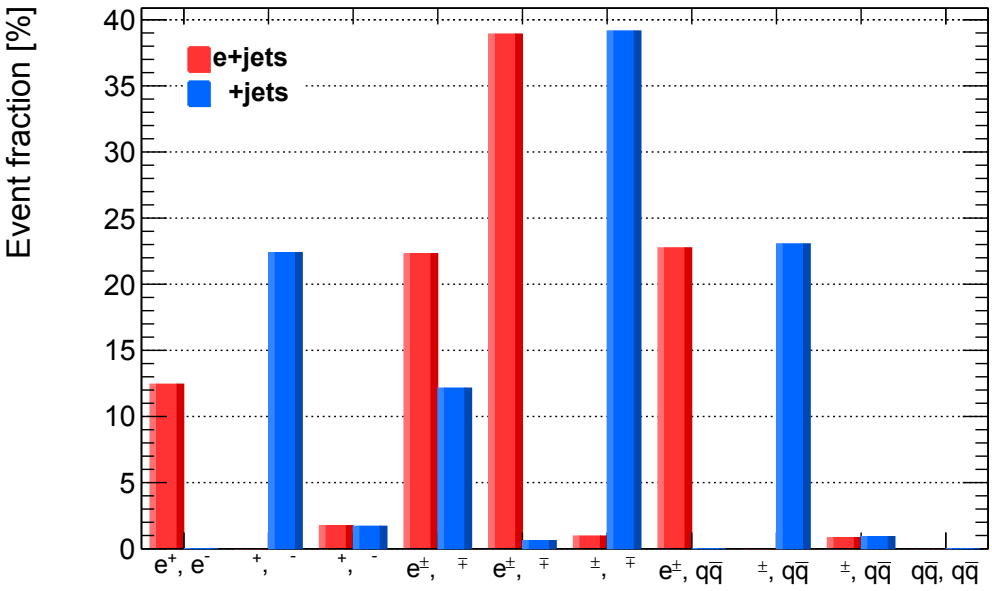


Figure C.4.: Event fractions of $t\bar{t}$ decay channels in the signal region without the di-lepton veto as predicted by the simulation. The electron and muon selections are considered separately. Most of the selected events originate from di-leptonic decays.

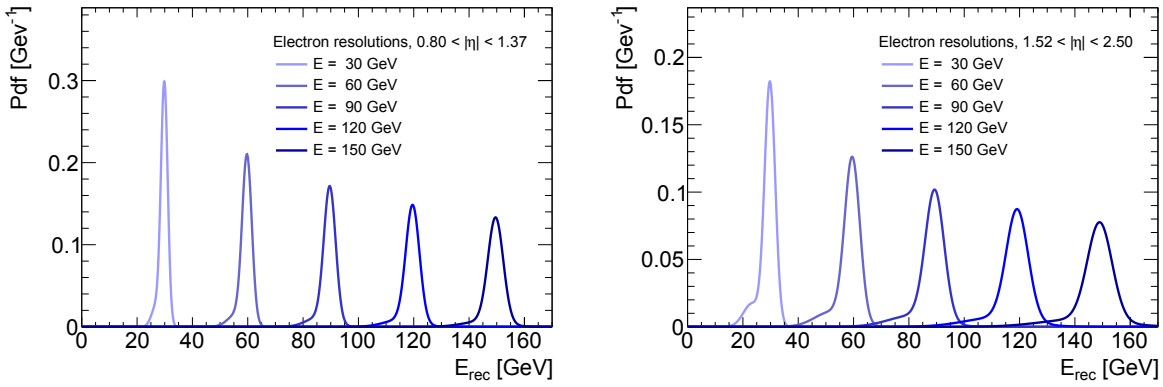
C.2. Di-lepton veto cut

A key element of the measurement presented in this thesis is the di-lepton veto, which is part of the event selection defining the signal region [170]. In order to reduce the number of di-leptonic top-quark pair production background events, a dedicated, loosened lepton definition is used as explained in Sec. 3.1.3. This allows one to reduce the number of these background events with minimal impact on the signal event yield.

Figure C.4 presents the event fractions of each decay channel of top-quark pair production as predicted by the simulation in the signal region without the di-lepton veto in case of the electron and muon selection, respectively. About 75% of the selected events come with two leptonic top-quark decays, where one of the charged leptons is not reconstructed given the tight selection criteria. Thus the di-lepton veto with its loosened lepton definition allows for a considerable reduction of the number of background events. This is shown in Tab. C.4, which provides the event yields of each process in the signal region with and without the di-lepton veto. The electron and muon event selections are combined. The signal event fraction is clearly increased due to this veto.

Process	without $\ell\ell$ veto	with $\ell\ell$ veto
s-channel	580 ± 20	570 ± 20
t-channel	1230 ± 40	1210 ± 40
Wt	500 ± 30	410 ± 30
$t\bar{t}$	11700 ± 100	7970 ± 90
W+jets	4700 ± 200	4700 ± 200
Z + jets & di-boson	570 ± 30	470 ± 30
Multi-jet	440 ± 20	370 ± 20
Total expectation	19720 ± 210	15740 ± 200
Data	19557	15556
S/B [%]	3.0	3.8

Table C.4.: Event yields with and without the di-lepton veto in the signal region. Each process is scaled according to the results of a dedicated fit using the m_{T}^{W} distribution. Only statistical uncertainties are provided. The bottom row shows the signal-to-background ratios, S/B. With the help of the di-lepton veto, the number of top-quark pair production background events is reduced, so that the signal event fraction is increased.



(a) (b)

Figure C.5.: Electron energy resolution functions. The probability densities of reconstructed energies given different particle energies are shown, where (a) and (b) refer to different ranges of pseudorapidity.

C.3. Resolution functions used for TELLs

The figures C.5 to C.7 show resolution functions of electrons, muons and b-jets which are used in the context of the TELL computations. They complement the discussion of Sec. 3.2.4 by referring to further regions of pseudorapidity. Clearly, the widths of the resolutions increase as the pseudorapidity increases [177].

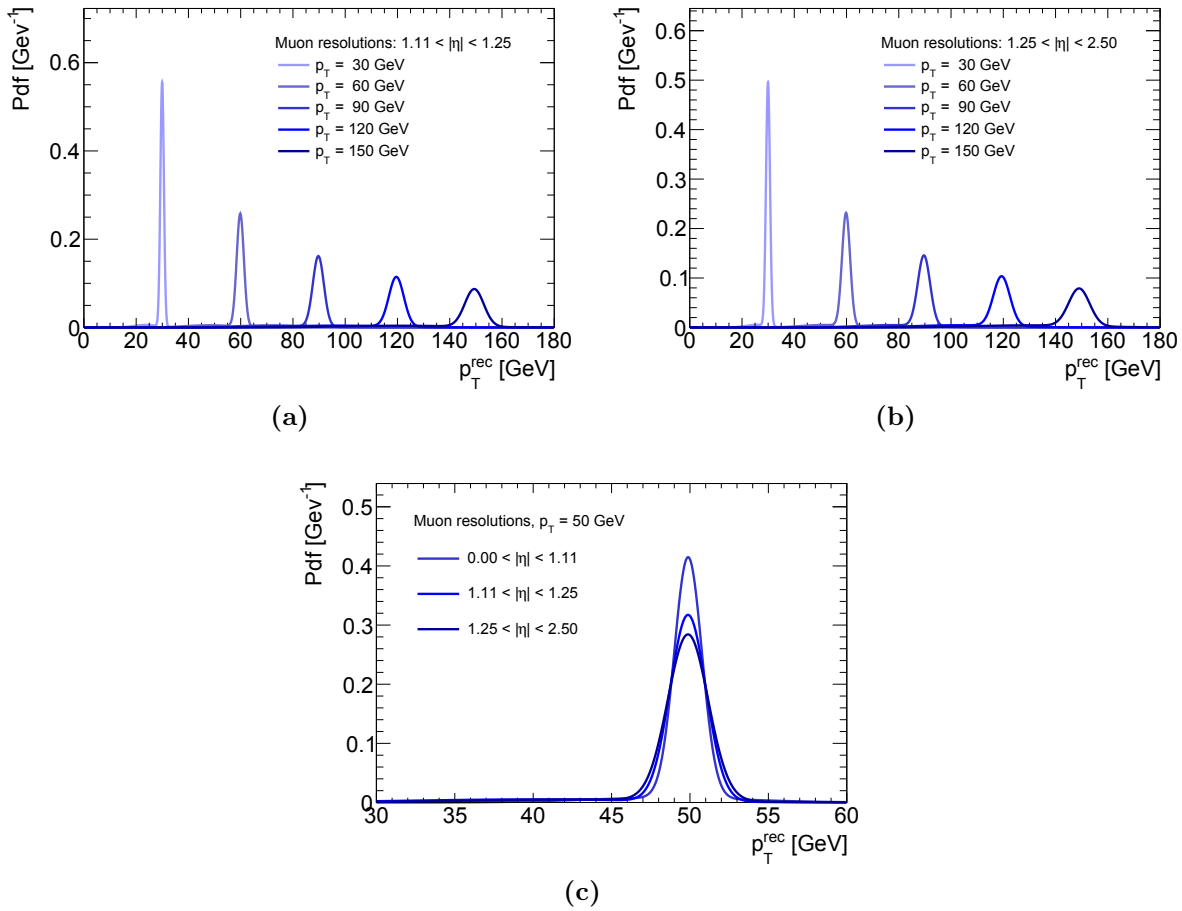


Figure C.6.: Muon p_{T} resolution functions. In (a) and (b) the probability density of reconstructed transverse momenta given different transverse momenta of the muon are shown. The two figures refer to different ranges of pseudorapidity. In (c) the resolution is shown for all pseudorapidity regions at a given transverse momentum of the muon.

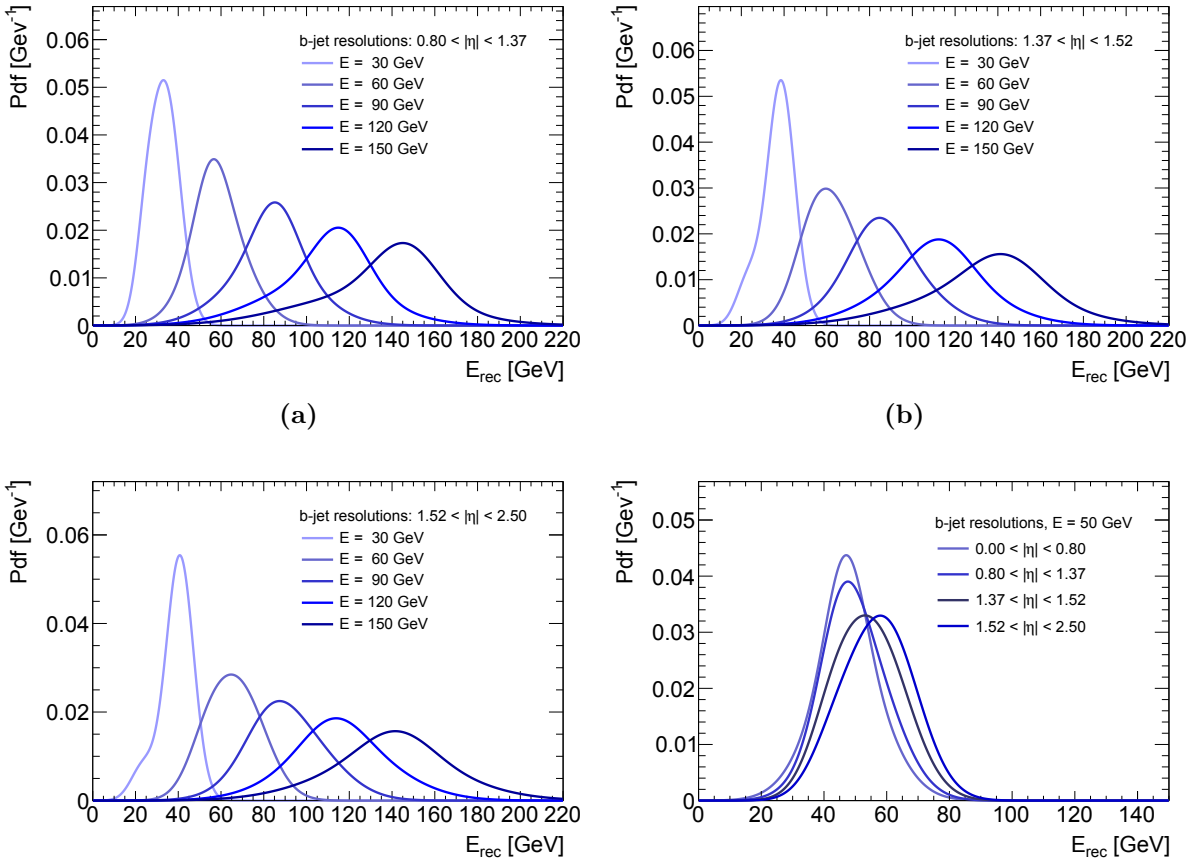


Figure C.7.: *b*-jet energy resolution functions. In (a) to (c) the probability densities of reconstructed energies given different parton energies are shown. Each of these figures refers to a different range of pseudorapidity. In (d) the resolution is shown for all pseudorapidity regions at a given *b*-quark energy.

C.4. Binning and modelling of the TELL discriminant

In the original histogram of the signal probability distribution $P(S|X)$ shown in Fig. 3.15, the single leftmost bin accumulates the background contributions. The signal contribution, on the other hand, is spread over many bins and there are only a few entries per bin. This binning is inappropriate with respect to the desired measurement, which makes use of a binned maximum likelihood fit. In order to obtain an appropriate histogram, a rebinning algorithm is applied. It constitutes a monotonic function of the original discriminant $P(S|X)$. The algorithm has already been applied in another measurement performed by the ATLAS collaboration [211].

First of all, a histogram of the distribution of $P(S|X)$ is created with many bins. Starting from $P(S|X) = 1$, bins are subsequently merged with the help of the following function of bin indices i, j:

$$Z(i,\, j) = z_S \frac{n_S(i,\, j)}{N_S} + z_B \frac{n_B(i,\, j)}{N_B} \;, \tag{C.1}$$

where

- $n_S(i,\, j)$, $n_B(i,\, j)$ are the numbers of signal and background events in the bins with indices between i and j, respectively;
- N_S, N_B are the numbers of all signal and background events, respectively.

Furthermore, z_S and z_B are parameters which can be adjusted in order to optimize the binning. The algorithm works as follows:

1. Starting on the right hand side of the distribution where $P(S|X) = 1$, $j = k_{\max}$, the range of bins in question is increased, so that one bin is added after the other, $(i,\, j) \to (i - 1,\, j)$.
2. At each step, $Z(i,\, j)$ is calculated.
3. As soon as $Z(i',\, j) > 1$, the bins of the interval $(i',\, j)$ are merged, unless the relative statistical uncertainty of the background contribution is greater than 10%, i. e. $\sqrt{n_B(i',\, j)}/n_B(i',\, j) > 0.1$.
4. These three steps are repeated, always starting from the last bin which has not yet been merged, until all bins are merged ($i' = 1$).

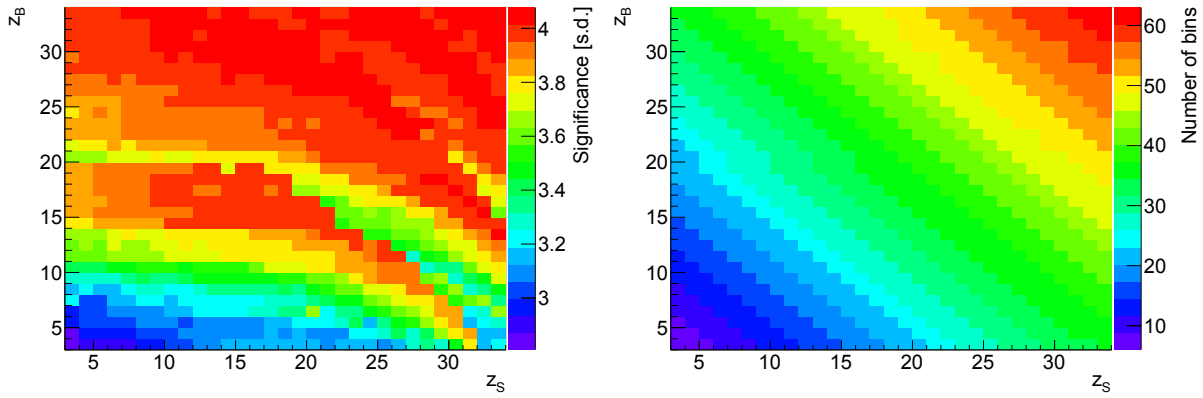
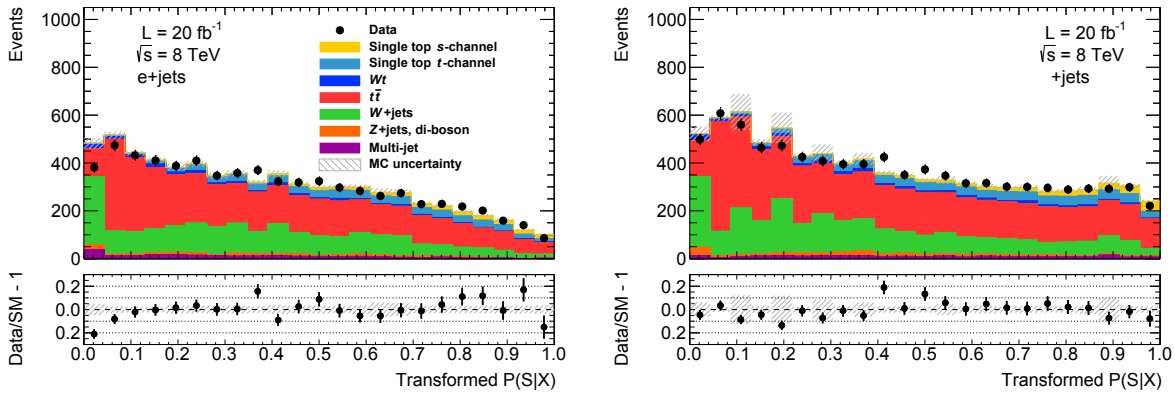


Figure C.8.: Study of the binning procedure of the discriminant. The expected signal significance is determined for several choices of the two parameters z_s and z_B (Eq. C.1). All uncertainties are included in this context. There is a broad range of parameters which give similar results (a). The number of bins grows with z_S and z_B (b).

Different values of z_S and z_B are tested with regard to the expected signal significance of the search for s-channel single top-quark production. In order to be able to test many different configurations efficiently, asymptotic formulae are used for the distribution of the profile likelihood ratio test statistic $\tilde{q}_0$ (Eq. 3.46). All systematic uncertainties are taken into account. As a result, there is a broad range of values of z_S and z_B, or, equivalently, a broad range of binnings which give similar results. Figure C.8 shows the expected signal significance and the number of bins for different choices of the parameters. The aim of the procedure is to find a binning which provides a high signal significance, given a relatively small number of bins in order to reduce the impact of possible fluctuations in the context of the evaluation of systematic uncertainties and their impact on the shape of the final discriminant.

Figure C.8 refers to the discriminant $P(S|X)$ constructed with the help of a-priori process probabilities $P(H_i)$ (Eq. 3.19) obtained in the context of the present analysis. They differ from the a-priori probabilities used in a previous analysis iteration published by the ATLAS collaboration where different reference cross sections of single top-quark production are used, as discussed in Sec. 3.2.6. In order to simplify comparisons with the latter analysis, the discriminant used in this thesis makes use of the same a-priori probabilities and uses the same binning. However, the differences between the two discriminants are small. In the previous analysis the parameters $z_S = 8$ and $z_B = 15$



(a) Electron selection **(b)** Muon selection

Figure C.9.: TELL discriminant distribution in the signal region for the electron (a) and muon (b) selection. All samples are scaled according to the corresponding control distribution fit results.

are chosen. Finally, a constant width is assigned to the resulting bins. This is a necessary constraint of the HistFactory tool which is used in the context of the statistical evaluation [195].

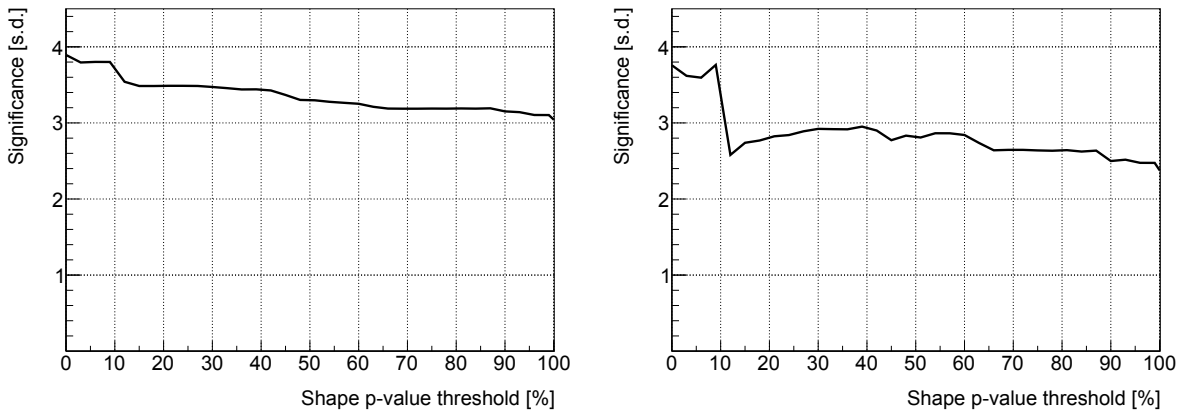
Subsequent to the binning of the discriminant, the goodness of its modelling by means of the employed predictions needs to be investigated. In Sec. 3.2.6 a proper modelling of the discriminant is shown in case of the W+jets and $t\bar{t}$ regions. Figure C.9 shows the discriminant distribution in the signal region in case of the electron and muon selection, respectively. The normalisation of each scattering process is scaled according to the control distribution fit results discussed in Sec. 3.1.4. Overall, a good agreement between the data and the expectation is achieved. The only exception is given by the leftmost bin in the electron case, Fig. C.9a. However, only statistical uncertainties are presented here. In particular, the large normalisation uncertainty of the W+jets background, which dominates this bin, comes in addition. Besides, the good modelling of the various considered kinematic distributions, single TELL distributions and TELL discriminants in the two control regions create confidence in the employed predictions. Thus it is concluded that the TELL discriminant is modelled properly and only its leftmost bin is left out in the context of the measurement.

C.5. W+jets modelling study

Discriminant shape variations due to systematic uncertainties are taken into account in the context of the statistical model if the p-value resulting from a χ^2 test comparing nominal and systematically varied distributions is less than 5%, as discussed in Sec. 3.3. The impact of this choice on the measurement is estimated by dedicated determinations of the signal significance given different p-value thresholds. The results are presented in Fig. C.10 in case of the expected and observed significance. Asymptotic formulae are used to determine the distribution of the employed test statistic $\tilde{q}_0$ (Eq. 3.46). As a result, the expected signal significance slightly decreases if more shape variations are considered. This behaviour is expected since these variations are also subject to statistical fluctuations which smear the discriminant distributions. The observed significance shows a similar behaviour except for the case of the W+jets modelling uncertainty. Apart from this source of uncertainty, these results show that the dependency of the measurement on the choice of the p-value threshold is small. The W+jets modelling, on the other hand, needs to be studied in more detail.

First of all it needs to be noticed that the difference of the discriminant shape related to the W+jets event simulation based on the Sherpa generator on the one hand, and on the Alpgen generator interfaced with Pythia 6 on the other hand, is insignificant as the p-value of the corresponding χ^2 test amounts to 10%. Still, differences of measurement results related to the choice between these event generators are possible due to the relatively low amount of available simulation statistics (Tab. A.1 and A.2). Figure C.11 shows the discriminant distribution as predicted by the two event generators.

In the following a study is presented where the Sherpa W+jets prediction is replaced by the one based on Alpgen + Pythia 6. The fit of this alternative model is performed in the same way as in case of the baseline model. Asymptotic formulae are used to determine the distributions of the relevant test statistics in this context [199]. The expected signal significance and cross section uncertainty based on the alternative model are equal to the results based on the baseline model, namely 3.8 standard deviations and 30%, respectively. In case of the observed results, deviations occur. Figure C.12 shows the discriminant distribution resulting from a fit of the model to the measured data using the Alpgen + Pythia 6 W+jets prediction. Again, the model fits the data. The fitted signal strength amounts to $\hat{\mu}^{\text{observed}}_{\text{Alpgen}} = 0.57$, which is less than the baseline result $\hat{\mu}^{\text{observed}}_{\text{Sherpa}} = 1.02$. The signal significance equals 2.1 instead of 3.6 standard deviations,



(a) Expectation **(b)** Measurement

Figure C.10.: Impact of discriminant shape variations on the signal significance. Variations of the discriminant shape due to systematic uncertainties are taken into account if the p-value of a corresponding χ^2 test comparing the distributions in question is below a defined threshold. The signal significance is determined as a function of this threshold in case of the Asimov data (a) and in case of the measured data (b). The impact of the choice of the threshold is small except for the inclusion of the W+jets modelling uncertainty in case of the measured data. Thus a dedicated study is performed in this respect.

while the cross section uncertainty amounts to 51% instead of 31%. The two confidence intervals overlap each other.

In order to investigate the difference of the observed results, the two models are compared with each other explicitly. The baseline model using the SHERPA W+jets prediction is used to generate pseudo-data sets of the discriminant distribution. Subsequent to each pseudo-data generation, the baseline and the alternative model are fit to this pseudo-data, respectively. The signal strengths resulting from these two fits are compared with each other, so that the expected difference of signal strengths resulting from fits of the two models to data can be estimated. The result is presented in Fig. C.13. Indeed, the signal strength derived from the fit of the model using the ALPGEN + PYTHIA 6 prediction is considerably smaller than the signal strength resulting from fits of the baseline model. This difference explains the observed deviation of the two signal strengths from each other. In particular, it is concluded that this deviation does not call the validity of the SHERPA W+jets prediction into question. Furthermore, this prediction comes with smaller statistical uncertainties. Thus it allows for a more precise measurement compared to the ALPGEN + PYTHIA 6 prediction. Given the insignificance of the difference of the two discriminant distribution shapes mentioned above, the explanation of the difference

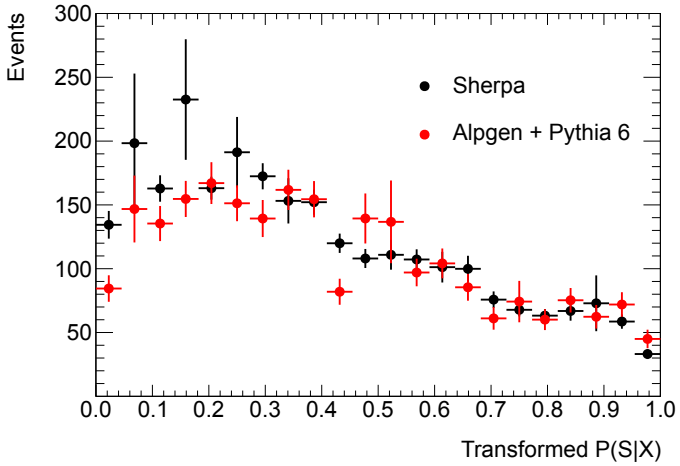


Figure C.11.: Discriminant distribution of W+jets events in the signal region predicted by different MC event generators.

of the observed fit results determined by means of the pseudo-experiments and the higher statistical precision of the SHERPA W+jets prediction, the latter is chosen in the context of the measurement. The uncertainty due to the modelling of W+jets events enters as an overall rate uncertainty, comparing the total event yields given by the two predictions.

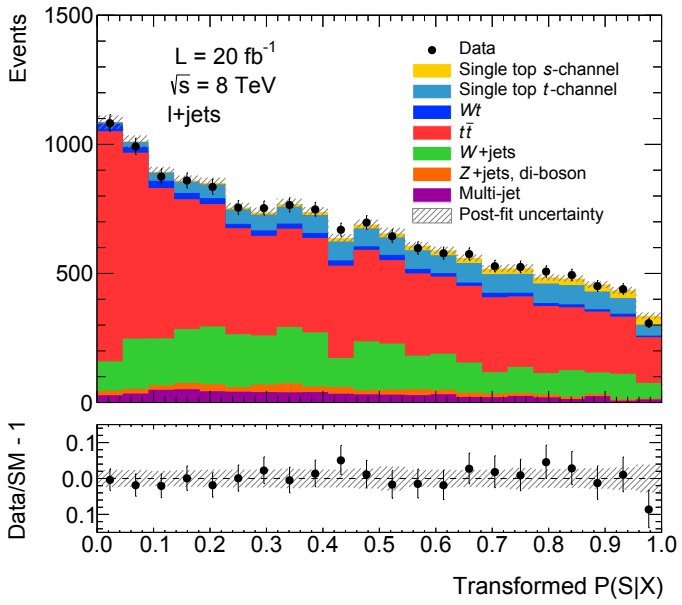


Figure C.12.: Fit of the TELL discriminant model to the data using the ALPGEN + PYTHIA 6 W+jets event generation.

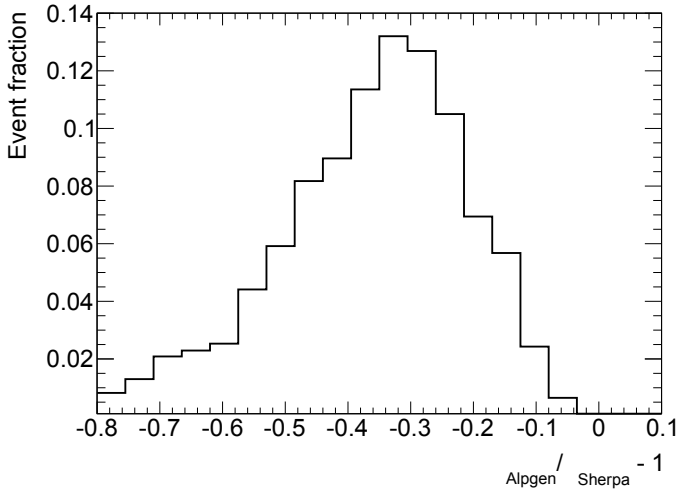
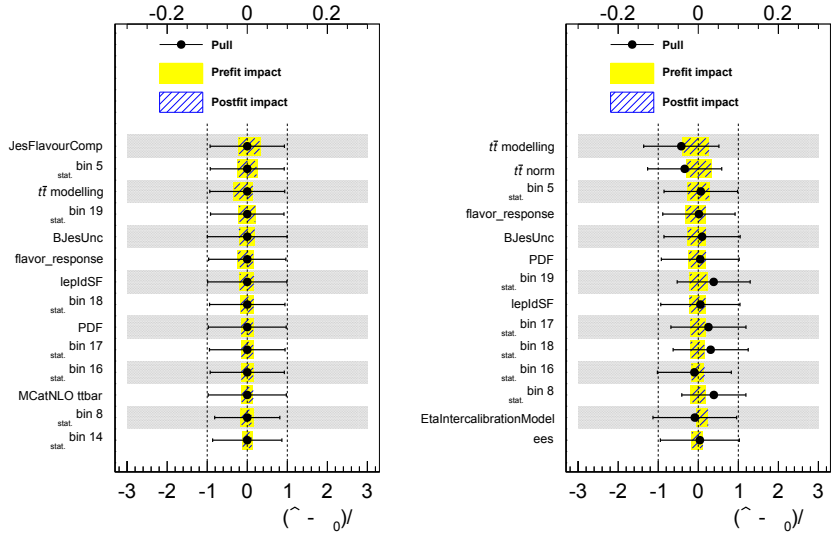


Figure C.13.: Comparison of fitted signal strengths using different W+jets event generators. Pseudo-data sets are generated using the model based on the SHERPA W+jets event generation. Subsequently, this model as well as the alternative model using the ALPGEN + PYTHIA 6 W+jets prediction is fit to this pseudo-data. The signal strengths resulting from fits of the latter model are relatively small. This finding matches the result of the fits to the oberved data, where the difference amounts to $\mu^{\text{observed}}_{\text{Alpgen}}/\mu^{\text{observed}}_{\text{Sherpa}} - 1 = -0.44$.

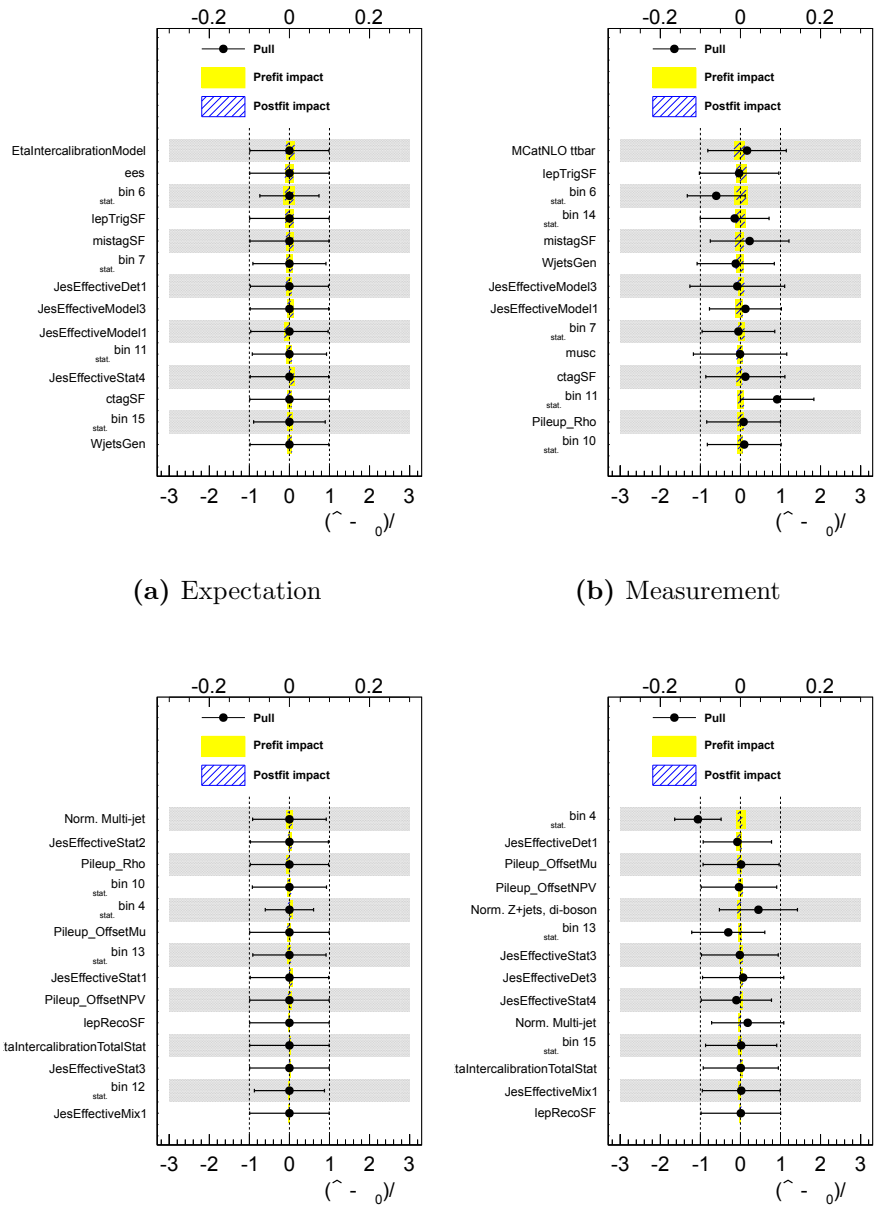


(a) Expectation **(b)** Measurement

Figure C.14.: Nuisance parameter pulls, extended. The fit results are based on the Asimov data (a) and on the measured data (b). The pulls provide the MLEs of nuisance parameters together with their fit uncertainties. In addition, the impact of nuisance parameters on the signal strength is provided. The observed results are similar to the expectation.

C.6. Fit results of the measurement

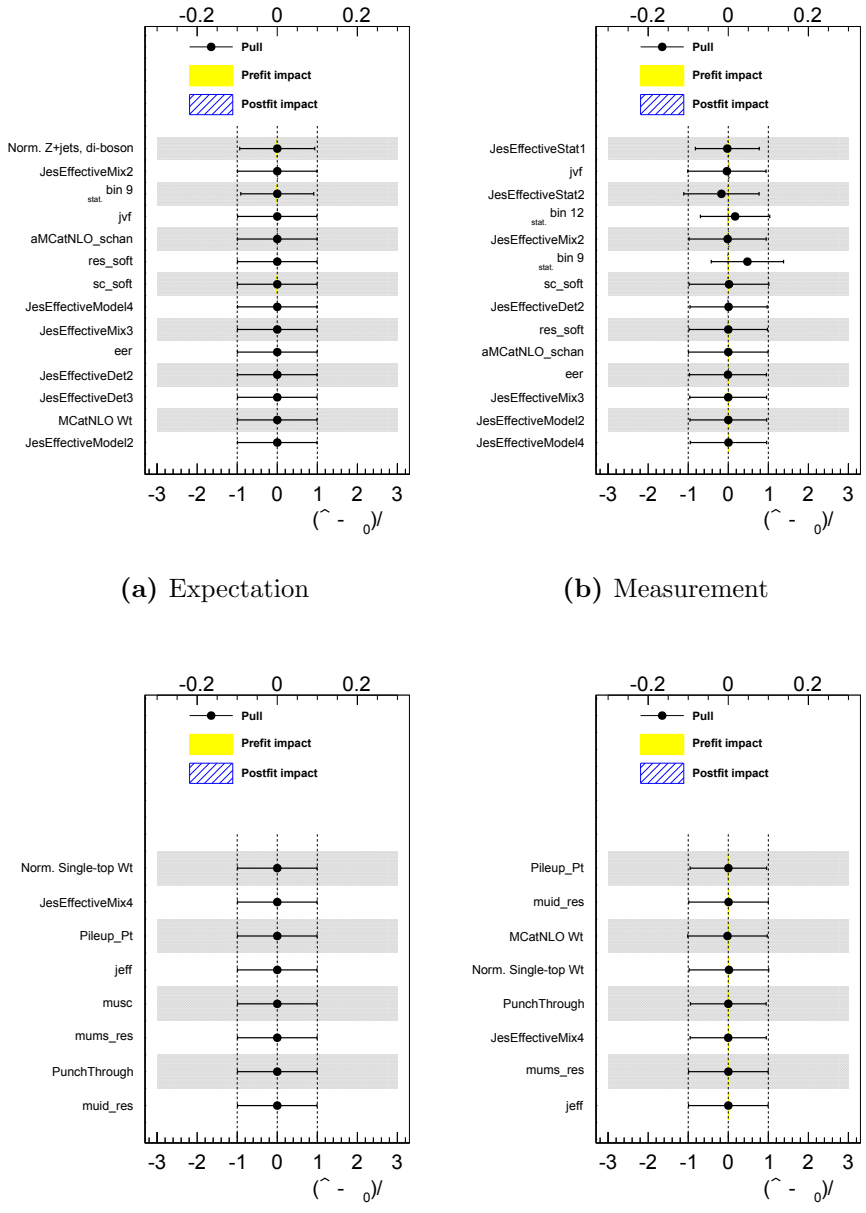
Figure C.14, C.15 and C.16 show the results of fits of the model to the Asimov data and to the measured data. For each nuisance parameter the MLE, the uncertainty resulting from the fit and the impact on the signal strength μ before and after the fit are shown. This impact is defined as described in Sec. 3.4.1. The nuisance parameters are ranked according to their impact on μ. These results complete the information provided in Fig. 3.23, where only the nuisance parameters with large impacts on μ are shown. Again, the deviations of the observed MLEs from the expectation α_0 are small compared to the uncertainty $\Delta\alpha$ given by auxiliary knowledge except for some of the nuisance parameters associated with statistical uncertainties of the predictions and for the theoretical normalisation of the Z+jets and di-boson background. Furthermore, only a few nuisance parameters are further constrained by the fit to the measured data, in particular the nuisance parameters which are associated with statistical uncertainties. In summary, these are reasonable fit results.



(a) Expectation **(b)** Measurement

(c) Expectation **(d)** Measurement

Figure C.15.: Nuisance parameter pulls, extended. The fit results are based on the Asimov data (a), (c) and on the measured data (b), (d). The pulls provide the MLEs of nuisance parameters together with their fit uncertainties. In addition, the impact of nuisance parameters on the signal strength is provided. The observed results are similar to the expectation.



(a) Expectation **(b)** Measurement

(c) Expectation **(d)** Measurement

Figure C.16.: Nuisance parameter pulls, extended. The fit results are based on the Asimov data (a), (c) and on the measured data (b), (d). The pulls provide the MLEs of nuisance parameters together with their fit uncertainties. In addition, the impact of nuisance parameters on the signal strength is provided. The observed results are similar to the expectation.

Appendix D.

Tell phase space generation

While some parts of the TELL computations are collected from external sources, the integration over phase spaces is implemented specifically. Two kinds of phase space integrations are performed. First there are the total phase space integrations necessary for proper likelihood normalizations. Well established techniques are used like decompositions of phase spaces into multiple two-particle ones and these shall not be repeated here. Secondly, there are the event likelihood phase space integrations in which δ-functional transfer functions must first be integrated analytically. As the efficiency and accuracy of the analysis depend on these integrations, they are explained in the following.

The integration measure to be used in the likelihood computations must be given in terms of random numbers, denoted by $r \in [0,\, 1]$ in the following, in order to allow for numerical integrations by means of MC techniques. The momenta of final state particles are denoted by $p = E\,(\sin\vartheta\cos\varphi,\, \sin\vartheta\sin\varphi,\, \cos\vartheta)$. They are always assumed to be massless since even the b-quark mass, which is the largest final state particle mass in the present context, is much smaller than the scales of the processes in question, e.g. like the W boson mass or the top-quark mass. The integration measure must also involve the initial state parton momentum fractions $x_{1,2}$ of the proton momenta $P_{1,2} = \frac{\sqrt{s}}{2}(1,\, 0,\, 0,\, \pm 1)$ with $\sqrt{s}$ being the hadronic centre-of-mass energy. Through the transfer functions, reconstructed momentum components enter, which are denoted by the superscript "rec". The phase space measure for n particles p matched to a reconstructed object and m unmatched particles $\tilde{p}$ with $P = \sum_{i=1}^{n} p_i + \sum_{i=1}^{m} \tilde{p}_i$ denoting the momentum

sum of these particles is given by

$$\begin{aligned}\int \mathrm{d}\Phi_{n,m} = & \int \mathrm{d}x_1 \mathrm{d}x_2 \, (2\pi)^4 \, \delta^{(4)} \left(x_1 P_1 + x_2 P_2 - P\right) \\ & \cdot \prod_{i=1}^{n} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i} \delta\left(\cos\vartheta_i - \cos\vartheta_i^{\mathrm{rec}}\right) \delta\left(\varphi_i - \varphi_i^{\mathrm{rec}}\right) \\ & \cdot \prod_{j=1}^{m} \frac{\mathrm{d}^3 \tilde{p}_j}{(2\pi)^3 2\tilde{E}_j} \ . \end{aligned} \tag{D.1}$$

D.1. Generic phase space

The generic phase space generation is designed for final states with at least one neutrino plus any number of leptons and jets. Being the most straightforward implementation, it is not fast enough for the present analysis, but it serves as an important cross check for the W+jets and single top-quark phase space generations. Also it is used to check for proper transfer function normalisations. Let one neutrino be denoted by $p_\nu = \tilde{p}_m$. The mapping of random numbers to momenta is defined by

$$\begin{aligned} E_i &= 2\, E_i^{\mathrm{rec}}\, r_{E_i} \\ \tilde{E}_i &= \frac{\sqrt{s}}{2}\, r_{\tilde{E}_i} \\ \cos\tilde{\vartheta}_i &= 2\, r_{\cos\tilde{\vartheta}_i} - 1 \\ \tilde{\varphi}_i &= 2\pi\, r_{\tilde{\varphi}_i} \\ p_\nu^z &= \sqrt{s} \left(r_{\nu_z} - \frac{1}{2} \right) \ . \end{aligned} \tag{D.2}$$

It is sufficient to restrict the energies of matched particles to $\tilde{E}_i^{\mathrm{rec}} < 2\, \tilde{E}_i^{\mathrm{rec}}$ because the energy transfer functions vanish beyond this region. The directions of matched particles equal those of their reconstructed counterparts due to the assumed flawless angular resolution,

$$\begin{aligned} \cos\vartheta_i &= \cos\vartheta_i^{\mathrm{rec}} \ , \\ \varphi_i &= \varphi_i^{\mathrm{rec}} \ . \end{aligned} \tag{D.3}$$

The x and y-components of the δ-functionals which express momentum conservation in Eq. D.1 are solved by integrating the transverse neutrino momentum components p_ν^x, p_ν^y,

$$p_\nu^{x,y} = -\sum_{i=1}^{n} p_i^{x,y} - \sum_{i=1}^{m-1} \tilde{p}_i^{x,y} \ , \tag{D.4}$$

while the energy and z-components are solved by integrating x_1 and x_2,

$$x_1 = \frac{P^0 + P^3}{\sqrt{s}} \ , \qquad x_2 = \frac{P^0 - P^3}{\sqrt{s}} \ ,$$

$$\int_0^1\int \mathrm{d}x_1\mathrm{d}x_2\, \delta\left(x_1 P_1^0 + x_2 P_2^0 - P^0\right) \delta\left(x_1 P_1^3 + x_2 P_2^3 - P^3\right) = \frac{2}{s}\, \chi\left(0 < x_{1,2} < 1\right) \ , \tag{D.5}$$

where χ is the characteristic function,

$$\chi\left(\text{expression}\right) = \begin{cases} 1, & \text{if the expression is true} \\ 0, & \text{otherwise} \end{cases} \ . \tag{D.6}$$

This results in

$$\int \mathrm{d}\Phi_{n,m} = \frac{\sqrt{s}^{\,m-2}}{(2\pi)^{3n+2m-5} 2^{m-1}} \prod_{i=1}^{n} E_i^{\mathrm{rec}} \int \mathrm{d}^{n+3m-2} r\, \frac{\prod_{i=1}^{n} E_i \prod_{j=1}^{m-1} \tilde{E}_j}{E_\nu} \chi\left(0 < x_{1,2} < 1\right) \ . \tag{D.7}$$

D.2. W+jets production

The W+jets phase space generation is designed for events with n partons p matched to jets, m unmatched partons $\tilde{p}$, one charged lepton p_ℓ and one neutrino p_ν, where the latter two come from a W boson decay. The charged lepton must be matched to a reconstructed one. The mapping of random numbers to momenta is realised in the same way as it is done for the generic phase space generation (Eq. D.2) except for the longitudinal neutrino

momentum p_ν^z which is determined from the W boson mass:

$$\int \mathrm{d}p_\nu^z = \int \mathrm{d}p_\nu^z \mathrm{d}m_W^2 \, \delta\left(m_W^2 - 2p_\ell \cdot p_\nu\right) \tag{D.8}$$

$$= \int \mathrm{d}m_W^2 \mathrm{d}p_\nu^z \left(\frac{1}{\left|\frac{\mathrm{d}m_W^2}{\mathrm{d}p_\nu^z}(p_{\nu_a}^z)\right|} \delta\left(p_\nu^z - p_{\nu_a}^z\right) + \frac{1}{\left|\frac{\mathrm{d}m_W^2}{\mathrm{d}p_\nu^z}(p_{\nu_b}^z)\right|} \delta\left(p_\nu^z - p_{\nu_b}^z\right) \right) . \tag{D.9}$$

The equation $2p_\ell \cdot p_\nu = m_W^2$ extracted from Eq. D.8 has two solutions used in Eq. D.9,

$$p_{\nu_{a,b}}^z = \frac{1}{2\left(p_\ell^{\mathrm{T}}\right)^2} \left(m_W^2 p_\ell^{\mathrm{z}} + 2p_\ell^{\mathrm{z}} \left(\vec{p}_\ell^{\,\mathrm{T}} \cdot \vec{p}_\nu^{\,\mathrm{T}}\right) \pm 2E_\ell \sqrt{D_\nu} \right) ,$$

$$D_\nu \coloneqq \left(\vec{p}_\ell^{\,\mathrm{T}} \cdot \vec{p}_\nu^{\,\mathrm{T}}\right)^2 + m_W^2 \left(\vec{p}_\ell^{\,\mathrm{T}} \cdot \vec{p}_\nu^{\,\mathrm{T}}\right) + \left(\vec{p}_\nu^{\,\mathrm{T}}\right)^2 \left((p_\ell^z)^2 - E_\ell^2\right) + \frac{1}{4}\left(m_W^2\right)^2 , \tag{D.10}$$

where $\vec{p_{\mathrm{T}}} = (p_x, p_y)$ denotes the transverse momenta. The integration is restricted to the part of random number space for which the solutions $p_{\nu_{a,b}}^z$ are real. This condition is expressed by the corresponding characteristic function $\chi(p_{\nu_{a,b}}^z \in \mathbb{R}) = \chi(D_\nu \geq 0)$.

In order to speed up the likelihood computation, the W boson mass is generated in a way which counteracts the sharp W boson propagator which appears in the scattering amplitudes of the processes in question.

$$m_W^2 = M_W \Gamma_W \tan\left(\pi\left(r_{m_W^2} - \frac{1}{2}\right)\right) + M_W^2$$

$$\frac{\mathrm{d}m_W^2}{\mathrm{d}r_{m_W^2}} = \frac{\pi}{M_W \Gamma_W}\left((m_W^2 - M_W^2)^2 + M_W^2 \Gamma_W^2\right) \tag{D.11}$$

M_W and Γ_W denote the pole mass and the decay width of the W boson, respectively. Altogether the phase space measure takes the following form

$$\int \mathrm{d}\Phi_{n,m} = \frac{\sqrt{s}^{\,m-2}}{(2\pi)^{3n+2m+2} 2^m} E_\ell^{\mathrm{rec}} \prod_{i=1}^{n} E_i^{\mathrm{rec}} \int \mathrm{d}^{n+3m+2} r \sum_{p_\nu^z \in \{p_{\nu_a}^z, p_{\nu_b}^z\}} E_\ell \frac{\prod_{i=1}^{n} E_i \prod_{j=1}^{m} \tilde{E}_j}{E_\nu}$$

$$\cdot \frac{1}{\left|2E_\ell \frac{p_\nu^z}{E_\nu} - 2p_\ell^z\right|} \frac{\pi}{M_W \Gamma_W}\left((2p_\ell \cdot p_\nu - M_W^2)^2 + M_W^2 \Gamma_W^2\right)$$

$$\cdot \chi\left(0 < x_{1,2} < 1\right) \chi\left(D_\nu > 0\right) . \tag{D.12}$$

The sum over the solutions for p_ν^z is realized explicitly by performing the integral for both cases and summing up the two contributions.

D.3. Single top-quark production

Dedicated phase space generations are needed to perform sufficiently fast likelihood computations for single top-quark processes. Two different final states need to be considered. There are single top-quarks with one additional parton and with two additional partons. Starting with Eq. D.1, in both cases p_1, p_2, and p_3 denote the neutrino, the charged lepton and the b-quark coming from the top-quark decay, respectively, $p_1 = p_\nu$, $p_2 = p_\ell$, $p_3 = p_b$. In both cases the additional final state b-quark is denoted by p_4, $p_4 = p_{\bar{b}}$. The phase space integral of Eq. D.1 is transformed by inserting additional integrations which do not change the overall integral, similar to Eq. D.8. In this case an integration over the top-quark four-momentum is introduced [202]. The final state with two partons is discussed first. Subsequently, the three parton final state is treated accordingly.

Two final state partons

All final state particles except for the neutrino are assumed to be matched to a reconstructed object. Inserting the aforementioned top-quark four-momentum integration into Eq. D.1 gives

$$
\begin{aligned}
\int \mathrm{d}\Phi_{3,1} &= \int \mathrm{d}\Phi_{3,1} \int \frac{\mathrm{d}^4 p_t}{(2\pi)^4} \frac{\mathrm{d}m_t^2}{2\pi} (2\pi)^4 \, \delta^{(4)}(p_t - p_1 - p_2 - p_3) \chi(p_t^0 > 0) \, 2\pi \, \delta(m_t^2 - p_t^2) \\
&= \int \mathrm{d}x_1 \mathrm{d}x_2 \frac{\mathrm{d}m_t^2}{2\pi} \int \frac{\mathrm{d}^3 p_t}{(2\pi)^3 2E_t} \prod_{i=1}^{4} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i} \prod_{j=2}^{4} \delta\left(\cos\vartheta_j - \cos\vartheta_j^{\mathrm{rec}}\right) \delta\left(\varphi_j - \varphi_j^{\mathrm{rec}}\right) \\
&\quad \cdot (2\pi)^4 \delta^{(4)}\left(x_1 P_1 + x_2 P_2 - p_t - p_4\right) (2\pi)^4 \delta^{(4)}\left(p_t - p_1 - p_2 - p_3\right) \\
&= \int \mathrm{d}x_1 \mathrm{d}x_2 \frac{\mathrm{d}m_t^2}{2\pi} \int \frac{\mathrm{d}^3 p_t}{(2\pi)^3 2E_t} \frac{\mathrm{d}^3 p_1}{(2\pi)^3 2E_1} \prod_{i=2}^{4} \frac{\mathrm{d}E_i E_i}{(2\pi)^3 2} \\
&\quad \cdot (2\pi)^4 \delta\left((x_1 + x_2)\frac{\sqrt{s}}{2} - \sqrt{\vec{p}_t^{\,2} + m_t^2} - E_4\right) \delta^{(3)}\left((x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_t - \vec{p}_4\right) \\
&\quad \cdot (2\pi)^4 \delta(E_t - E_1 - E_2 - E_3) \delta^{(3)}\left(\vec{p}_t - \vec{p}_1 - \vec{p}_2 - \vec{p}_3\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{32(2\pi)^8} \int \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}m_t^2 \prod_{i=2}^{4} \mathrm{d}E_i E_i \frac{1}{E_1 E_t} \\
&\cdot \delta \left((x_1 + x_2)\frac{\sqrt{s}}{2} - \sqrt{\left| (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_4 \right|^2 + m_t^2} - E_4 \right) \\
&\cdot \delta \left(E_t - \left| (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_2 - \vec{p}_3 - \vec{p}_4 \right| - E_2 - E_3 \right)
\end{aligned} \tag{D.13}
$$

$\vec{e_z} = (0,\, 0,\, 1)$ denotes the z-direction. The delta-functionals at the end of Eq. D.13 can be transformed and integrated once their arguments are solved for E_4 and E_3, respectively. Rearranging and squaring the argument of the first delta-functional gives the solution E_4:

$$
E_4 = (x_1 + x_2)\frac{\sqrt{s}}{2} - \sqrt{\left| (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_4 \right|^2 + m_t^2} \,, \tag{D.14}
$$

$$
E_4 = \frac{x_1 x_2 \frac{s}{2} - \frac{m_t^2}{2}}{(x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_4} \,. \tag{D.15}
$$

In the following the usage of the top-quark energy is useful:

$$
\begin{aligned}
E_t &= \left| (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_4 \right|^2 \,, \\
E_t &= \sqrt{E_4^2 + (x_1 - x_2)^2 \frac{s}{4} - 2(x_1 - x_2)\frac{\sqrt{s}}{2} E_4 \cos\vartheta_4 + m_t^2} \,.
\end{aligned} \tag{D.16}
$$

Rearranging and squaring the argument of the second delta-functional in Eq. D.13 gives the solution E_3:

$$
\begin{aligned}
E_3 &= E_t - \left| (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_2 - \vec{p}_3 - \vec{p}_4 \right| - E_2 \,, \\
E_3 &= \frac{\frac{m_t^2}{2} - E_t E_2 - \vec{p}_2 \cdot \vec{p}_4 + (x_1 - x_2)\frac{\sqrt{s}}{2} E_2 \cos\vartheta_2}{E_t - E_2 + \vec{e}_3 \cdot (\vec{p}_2 + \vec{p}_4) - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_3} \,,
\end{aligned} \tag{D.17}
$$

where $\vec{e_3} = \vec{p_3}/E_3$ denotes the direction of particle p_3. Finally, p_1 results from momentum conservation,

$$
\begin{aligned}
\vec{p}_1 &= \vec{p}_t - \vec{p}_2 - \vec{p}_3 \,, \\
\vec{p}_1 &= (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_2 - \vec{p}_3 - \vec{p}_4 \,.
\end{aligned} \tag{D.18}
$$

As already discussed in the previous sections, the directions of matched particles equal the directions of the corresponding reconstructed objects,

$$\begin{aligned} \cos\vartheta_{2,3,4} &= \cos\vartheta_{2,3,4}^{\text{rec}} \ , \\ \varphi_{2,3,4} &= \varphi_{2,3,4}^{\text{rec}} \ . \end{aligned} \tag{D.19}$$

The solutions presented above are unique. This means that here only one integration is needed instead of a sum over multiple solutions as in the case of the $W+$jets phase space of Sec. D.2. The remaining integration variables depend on the random numbers as follows:

$$\begin{aligned} x_1 &= r_{x_1} \ , \\ x_2 &= r_{x_2} \ , \\ E_2 &= 2\, E_2^{\text{rec}}\, r_{E_2} + E^{\text{min}} \ , \\ m_t^2 &= 2M_t^2\, r_{m_t^2} \ . \end{aligned} \tag{D.20}$$

E^{min} denotes a small minimum energy of the charged lepton momentum $p_2 = p_\ell$ on the order of 1 MeV, necessary for stable numerical integrations. $M_t = 172.5\,\text{GeV}$ denotes the mass for which the top-quark propagator is maximal. There are a couple of boundary conditions given by energy-momentum conservation. The transformations of Eq. D.20 with $r \in [0,\, 1]$ do not necessarily meet these conditions. The initial state energy must be large enough to allow for the production of a top-quark with a mass of m_t, the energies E_3 and E_4 given by Eq. D.17 and D.26 must be positive, and energy must be conserved. These boundary conditions are simply taken into account by characteristic functions. In practice the MC integration algorithm converges to the set of random numbers which meet these conditions. In principle the integration could be improved by including these conditions in the random number mapping.

Once the delta-functionals of Eq. D.13 are transformed and integrated, the resulting phase space measure is

$$\begin{aligned}\int \mathrm{d}\Phi_{3,1} = \frac{1}{8(2\pi)^8} M_t^2 E_2^{\mathrm{rec}} \int \mathrm{d}^4 r \, E_2 \chi(x_1 x_2 s > m_t^2) \chi(E_3 > 0) \chi(E_4 > 0) \\ \cdot \chi(\sqrt{x_1 x_2 s} = E_1 + E_2 + E_3 + E_4) \\ \cdot \frac{x_1 x_2 \frac{s}{2} - \frac{m_t^2}{2}}{\left((x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_4\right)^2} \\ \cdot \frac{\frac{m_t^2}{2} - E_t E_2 - \vec{p}_2 \cdot \vec{p}_4 + (x_1 - x_2)\frac{\sqrt{s}}{2} E_2 \cos\vartheta_2}{\left(E_t - E_2 + \vec{e}_3 \cdot (\vec{p}_2 + \vec{p}_4) - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_3\right)^2} \, . \end{aligned} \tag{D.21}$$

Three final state partons

The third final state parton is not matched to a jet in the likelihood computations relevant to the s-channel search presented in this thesis. It is denoted by $\tilde{p}_1$ in the following. The phase space is generated in a way similar to the case of two final state partons discussed above, only expanding the derivation by $\tilde{p}_1$:

$$\begin{aligned}\int \mathrm{d}\Phi_{3,2} &= \int \mathrm{d}\Phi_{3,2} \int \frac{\mathrm{d}^4 p_t}{(2\pi)^4} \frac{\mathrm{d}m_t^2}{2\pi} (2\pi)^4 \delta^{(4)}(p_t - p_1 - p_2 - p_3) \chi(p_t^0 > 0) \, 2\pi \, \delta(m_t^2 - p_t^2) \\ &= \int \mathrm{d}x_1 \mathrm{d}x_2 \frac{\mathrm{d}m_t^2}{2\pi} \int \frac{\mathrm{d}^3 p_t}{(2\pi)^3 2E_t} \prod_{i=1}^{4} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i} \frac{\mathrm{d}^3 \tilde{p}_1}{(2\pi)^3 2\tilde{E}_1} \\ &\quad \cdot \prod_{j=2}^{4} \delta\left(\cos\vartheta_j - \cos\vartheta_j^{\mathrm{rec}}\right) \delta\left(\varphi_j - \varphi_j^{\mathrm{rec}}\right) \\ &\quad \cdot (2\pi)^4 \delta^{(4)}\left(x_1 P_1 + x_2 P_2 - p_t - p_4 - \tilde{p}_1\right) (2\pi)^4 \delta^{(4)}\left(p_t - p_1 - p_2 - p_3\right) \, . \end{aligned} \tag{D.22}$$

Solving for the kinematics gives:

$$E_4 = \frac{x_1 x_2 \frac{s}{2} - \frac{m_t^2}{2} - \tilde{E}_1 \left((x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\tilde{\vartheta}_1\right)}{(x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_4 + \tilde{E}_1\left(\vec{e}_4 \cdot \tilde{\vec{e}}_1 - 1\right)} \, , \tag{D.23}$$

$$E_t = \sqrt{E_4^2 + \tilde{E}_1^2 + (x_1 - x_2)^2 \frac{s}{4} - 2(x_1 - x_2)\frac{\sqrt{s}}{2}\left(E_4 \cos\vartheta_4 + \tilde{E}_1 \cos\tilde{\vartheta}_1\right) + 2\vec{p}_4 \cdot \tilde{\vec{p}}_1 + m_t^2} \, , \tag{D.24}$$

$$E_3 = \frac{\frac{m_t^2}{2} - E_t E_2 - \vec{p}_2 \cdot (\vec{p}_4 + \tilde{\vec{p}}_1) + (x_1 - x_2)\frac{\sqrt{s}}{2} E_2 \cos\vartheta_2}{E_t - E_2 + \vec{e}_3 \cdot (\vec{p}_2 + \vec{p}_4 + \tilde{\vec{p}}_1) - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_3} \ , \tag{D.25}$$

$$\vec{p}_1 = (x_1 - x_2)\frac{\sqrt{s}}{2}\vec{e}_z - \vec{p}_2 - \vec{p}_3 - \vec{p}_4 - \tilde{\vec{p}}_1 \ . \tag{D.26}$$

The directions of matched particles are fixed as usual,

$$\begin{aligned} \cos\vartheta_{2,3,4} &= \cos\vartheta_{2,3,4}^{\mathrm{rec}} \ , \\ \varphi_{2,3,4} &= \varphi_{2,3,4}^{\mathrm{rec}} \ . \end{aligned} \tag{D.27}$$

The mapping of random numbers to momenta is done in the same way as for the case of two final state partons, while in addition the momentum $\tilde{\vec{p}}_1$ must be generated:

$$\begin{aligned} x_1 &= r_{x_1} \ , \\ x_2 &= r_{x_2} \ , \\ E_2 &= 2\, E_2^{\mathrm{rec}}\, r_{E_2} + E^{\min} \ , \\ m_t^2 &= 2 M_t^2\, r_{m_t^2} \ , \\ \tilde{E}_1 &= \frac{\sqrt{s}}{2} r_{\tilde{E}_1} + E^{\min} \ , \\ \cos\tilde{\vartheta}_1 &= 2\, r_{\cos\tilde{\vartheta}_1} - 1 \ , \\ \tilde{\varphi}_1 &= 2\pi\, r_{\tilde{\varphi}_1} \ . \end{aligned} \tag{D.28}$$

Finally, the resulting phase space measure is:

$$\begin{aligned} \int \mathrm{d}\Phi_{3,2} = \frac{1}{16(2\pi)^{10}} \sqrt{s} M_t^2 E_2^{\mathrm{rec}} \int \mathrm{d}^7 r\, \tilde{E}_1 E_2 \chi(x_1 x_2 s > m_t^2)\chi(E_3 > 0)\chi(E_4 > 0) \\ \cdot \chi(\sqrt{x_1 x_2 s} = E_1 + E_2 + E_3 + E_4 + \tilde{E}_1) \\ \cdot \frac{x_1 x_2 \frac{s}{2} - \frac{m_t^2}{2} - \tilde{E}_1 \left((x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\tilde{\vartheta}_1 \right)}{\left((x_1 + x_2)\frac{\sqrt{s}}{2} - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_4 + \tilde{E}_1 \left(\vec{e}_4 \cdot \tilde{\vec{e}}_1 - 1 \right) \right)^2} \\ \cdot \frac{\frac{m_t^2}{2} - E_t E_2 - \vec{p}_2 \cdot (\vec{p}_4 + \tilde{\vec{p}}_1) + (x_1 - x_2)\frac{\sqrt{s}}{2} E_2 \cos\vartheta_2}{\left(E_t - E_2 + \vec{e}_3 \cdot (\vec{p}_2 + \vec{p}_4 + \tilde{\vec{p}}_1) - (x_1 - x_2)\frac{\sqrt{s}}{2}\cos\vartheta_3 \right)^2} \ . \end{aligned} \tag{D.29}$$

D.4. Top-quark pair production

The top-quark pair production phase space is generated according to Ref. [212], whereas for the case at hand the decay products of one of the W bosons are not matched to jets and the two b-quarks are assumed to be massless. The notation of the following discussion refers to the case of single lepton $t\bar{t}$ events, while the phase space of di-lepton $t\bar{t}$ events is generated accordingly.

One of the top-quarks decays into a neutrino, a charged lepton and a b-quark, $p_{t_\ell} = p_\nu + p_\ell + p_{b_\ell}$, the other one into two lighter quarks and another b-quark, $p_{t_q} = p_q + p_{\bar{q}} + p_{b_q}$. Only the charged lepton p_ℓ and the b-quarks p_{b_ℓ}, p_{b_q} are matched to reconstructed objects. Following the notation introduced at the very beginning of this section, this corresponds to $p_1 = p_\ell$, $p_2 = p_{b_\ell}$, $p_3 = p_{b_q}$, $\tilde{p}_1 = p_\nu$, $\tilde{p}_2 = p_q$ and $\tilde{p}_3 = p_{\bar{q}}$. The integration variables are chosen to be the two top-quark masses squared, $m^2_{t_\ell} = p^2_{t_\ell}$ and $m^2_{t_q} = p^2_{t_q}$, the mass of the hadronically decaying W boson squared, $m^2_{W_q} = (p_q + p_{\bar{q}})^2$, the energy of one of the final state light quarks, $E_q = |p_q|$, the z-component of the momentum sum of the neutrino p_ν and the b-quark p_{b_ℓ}, $p^z_{b_\ell\nu} = (p_\nu + p_{b_\ell}) \cdot \vec{e}_z$ and the directions $\cos\vartheta, \phi$ of the unmatched final state light quarks p_q and $p_{\bar{q}}$. From these variables the involved momenta can be constructed as follows,

$$\begin{aligned} m^2_{W_q} &= 2E_q E_{\bar{q}} \left(1 - \cos\vartheta_{q,\bar{q}}\right) \ , \\ E_{\bar{q}} &= \frac{m^2_{W_q}}{2E_q \left(1 - \cos\vartheta_{q,\bar{q}}\right)} \ . \end{aligned} \tag{D.30}$$

For the derivation of p_{b_q}, the momentum sum of the light quarks, $p_{q\bar{q}} = p_q + p_{\bar{q}}$, is used:

$$\begin{aligned} m^2_{t_q} &= (p_{q\bar{q}} + p_{b_q})^2 = m^2_{W_q} + 2p_{q\bar{q}} \cdot p_{b_q} \ , \\ E_{b_q} &= \frac{1}{2} \frac{m^2_{t_q} - m^2_{W_q}}{E_{q\bar{q}} - \sqrt{E^2_{q\bar{q}} - m^2_{W_q}} \cos\vartheta_{b_q,q\bar{q}}} \ . \end{aligned} \tag{D.31}$$

In order to derive p_{b_ℓ}, the momentum sum of p_ν and p_{b_ℓ} is used. As the longitudinal component $p^z_{b_\ell\nu}$ is an integration variable, it follows directly from one of the random numbers. The transverse components follow from momentum conservation,

$$p^{x,y}_{b_\ell\nu} = -p^{x,y}_{t_q} - p^{x,y}_\ell = -p^{x,y}_q - p^{x,y}_{\bar{q}} - p^{x,y}_{b_q} - p^{x,y}_\ell \ , \tag{D.32}$$

and the energy results from energy conservation,

$$\begin{aligned} E_{b_\ell\nu} = E_{t_\ell} - E_\ell &= \sqrt{(p^x_{b_\ell\nu\ell})^2 + (p^y_{b_\ell\nu\ell})^2 + (p^z_{b_\ell\nu} + p^z_\ell)^2 + m^2_{t_\ell}} - E_\ell \\ &= \sqrt{(p^x_{b_q q\bar{q}})^2 + (p^y_{b_q q\bar{q}})^2 + (p^z_{b_\ell\nu} + p^z_\ell)^2 + m^2_{t_\ell}} - E_\ell \ . \end{aligned} \tag{D.33}$$

Now the magnitude of p_{b_ℓ} can be determined,

$$\begin{aligned} m^2_{t_\ell} &= (p_\ell + p_\nu + p_{b_\ell})^2 = 2p_\ell \cdot p_{b_\ell} + 2p_\ell \cdot p_\nu + 2p_{b_\ell} \cdot p_\nu \\ &= 2p_\ell \cdot p_{b_\ell\nu} + 2p_{b_\ell} \cdot p_{b_\ell\nu} \\ &= 2E_{b_\ell}(E_{b_\ell\nu} - |\vec{p}_{b_\ell\nu}| \cos\vartheta_{b_\ell,\, b_\ell\nu}) + 2p_\ell \cdot p_{b_\ell\nu} \ , \\ E_{b_\ell} &= \frac{1}{2} \frac{m^2_{t_\ell} - 2p_\ell \cdot p_{b_\ell\nu}}{E_{b_\ell\nu} - |\vec{p}_{b_\ell\nu}| \cos\vartheta_{b_\ell,\, b_\ell\nu}} \ . \end{aligned} \tag{D.34}$$

The longitudinal neutrino momentum can be determined from $p_{b_\ell\nu}$ and p_{b_ℓ}, whereas its transverse momentum as well as the initial state momenta result from energy-momentum conservation,

$$p^z_\nu = p^z_{b_\ell\nu} - p^z_{b_\ell} \ , \tag{D.35}$$

$$p^{x,\,y}_\nu = - \sum_{i \,\in\, \{\mathrm{b_q},\, \mathrm{q},\, \bar{\mathrm{q}},\, \mathrm{b}_\ell,\, \ell\}} p^z_i \ , \tag{D.36}$$

$$x_1 = \frac{1}{\sqrt{s}} \sum_{i \,\in\, \{\mathrm{b_q},\, \mathrm{q},\, \bar{\mathrm{q}},\, \mathrm{b}_\ell,\, \ell,\, \nu\}} (E_i + p^z_i) \ , \tag{D.37}$$

$$x_2 = \frac{1}{\sqrt{s}} \sum_{i \,\in\, \{\mathrm{b_q},\, \mathrm{q},\, \bar{\mathrm{q}},\, \mathrm{b}_\ell,\, \ell,\, \nu\}} (E_i - p^z_i) \ . \tag{D.38}$$

The Jacobian of the transformation of variables is given by

$$\mathrm{d}E_{\bar{q}} \mathrm{d}E_{b_\ell} \mathrm{d}E_{b_q} \mathrm{d}p^z_\nu = \left| \frac{\partial(m^2_{t_\ell},\, m^2_{t_q},\, m^2_{W_q},\, p^z_{b_\ell\nu})}{\partial(E_{\bar{q}},\, E_{b_\ell},\, E_{b_q},\, p^z_\nu)} \right|^{-1} \mathrm{d}m^2_{t_\ell} \mathrm{d}m^2_{t_q} \mathrm{d}m^2_{W_q} \mathrm{d}p^z_{b_\ell\nu} \ , \tag{D.39}$$

$$\frac{\partial(m^2_{t_\ell},\, m^2_{t_q},\, m^2_{W_q},\, p^z_{b_\ell\nu})}{\partial(E_{\bar{q}},\, E_{b_\ell},\, E_{b_q},\, p^z_\nu)} = \begin{vmatrix} \frac{\partial m^2_{t_\ell}}{\partial E_{\bar{q}}} & \frac{\partial m^2_{t_\ell}}{\partial E_{b_\ell}} & \frac{\partial m^2_{t_\ell}}{\partial E_{b_q}} & \frac{\partial m^2_{t_\ell}}{\partial p^z_\nu} \\ \frac{\partial m^2_{t_q}}{\partial E_{\bar{q}}} & \frac{\partial m^2_{t_q}}{\partial E_{b_\ell}} & \frac{\partial m^2_{t_q}}{\partial E_{b_q}} & \frac{\partial m^2_{t_q}}{\partial p^z_\nu} \\ \frac{\partial m^2_{W_q}}{\partial E_{\bar{q}}} & \frac{\partial m^2_{W_q}}{\partial E_{b_\ell}} & \frac{\partial m^2_{W_q}}{\partial E_{b_q}} & \frac{\partial m^2_{W_q}}{\partial p^z_\nu} \\ \frac{\partial p^z_{b_\ell\nu}}{\partial E_{\bar{q}}} & \frac{\partial p^z_{b_\ell\nu}}{\partial E_{b_\ell}} & \frac{\partial p^z_{b_\ell\nu}}{\partial E_{b_q}} & \frac{\partial p^z_{b_\ell\nu}}{\partial p^z_\nu} \end{vmatrix}$$

$$
\begin{aligned}
&= \begin{vmatrix}
0 & \frac{\partial m^2_{t_\ell}}{\partial E_{b_\ell}} & 0 & \frac{\partial m^2_{t_\ell}}{\partial p^z_\nu} \\
\frac{\partial m^2_{t_q}}{\partial E_{\bar{q}}} & 0 & \frac{\partial m^2_{t_q}}{\partial E_{b_q}} & 0 \\
\frac{\partial m^2_{W_q}}{\partial E_{\bar{q}}} & 0 & 0 & 0 \\
0 & \frac{\partial p^z_{b_\ell\nu}}{\partial E_{b_\ell}} & 0 & \frac{\partial p^z_{b_\ell\nu}}{\partial p^z_\nu}
\end{vmatrix} \\
&= \frac{\partial m^2_{W_q}}{\partial E_{\bar{q}}} \frac{\partial m^2_{t_q}}{\partial E_{b_q}} \left(\frac{\partial m^2_{t_\ell}}{\partial E_{b_\ell}} \frac{\partial p^z_{b_\ell\nu}}{\partial p^z_\nu} - \frac{\partial m^2_{t_\ell}}{\partial p^z_\nu} \frac{\partial p^z_{b_\ell\nu}}{\partial E_{b_\ell}} \right) \ . \qquad \text{(D.40)}
\end{aligned}
$$

The following derivatives are needed to evaluate Eq. D.40:

$$
\begin{aligned}
\frac{\partial p^{x,y}_\nu}{\partial E_{b_\ell}} &= \ - e^{x,y}_{b_\ell} \ , \\
\frac{\partial p^z_\nu}{\partial E_{b_\ell}} &= 0 \ , \\
\frac{\partial E_\nu}{\partial E_{b_\ell}} &= \ - \vec{e}^{\,\mathrm{T}}_\nu \cdot \vec{e}^{\,\mathrm{T}}_{b_\ell} \ . \qquad \text{(D.41)}
\end{aligned}
$$

Equation D.36 was used here. The vectors $\vec{e}_i = \vec{p}_i/E_i$ denote the directions of the momenta $\vec{p}_i$, whereas $\vec{e}^{\,\mathrm{T}}_i = (e^{\mathrm{x}}_i,\, e^{\mathrm{y}}_i,\, 0)$ denotes the corresponding transverse component.

The derivatives needed to build the Jacobian given by Eq. D.40 are

$$\begin{aligned}
\frac{\partial m_{W_q}^2}{\partial E_{\bar{q}}} &= \frac{\partial}{\partial E_{\bar{q}}} \left(2E_q E_{\bar{q}} \left(1 - \cos\vartheta_{q,\bar{q}}\right)\right) \\
&= 2E_q \left(1 - \cos\vartheta_{q,\bar{q}}\right) \ , && \text{(D.42)} \\
\frac{\partial m_{t_q}^2}{\partial E_{b_q}} &= \frac{\partial}{\partial E_{b_q}} \left(2p_q \cdot p_{\bar{q}} + 2p_q \cdot p_{b_q} + 2p_{\bar{q}} \cdot p_{b_q}\right) \\
&= 2E_q \left(1 - \cos\vartheta_{q,b_q}\right) + 2E_{\bar{q}} \left(1 - \cos\vartheta_{\bar{q},b_q}\right) \ , && \text{(D.43)} \\
\frac{\partial m_{t_\ell}^2}{\partial E_{b_\ell}} &= \frac{\partial}{\partial E_{b_\ell}} \left(2p_{b_\ell} \cdot p_\ell + 2p_{b_\ell} \cdot p_\nu + 2p_\ell \cdot p_\nu\right) \\
&= 2(E_\ell + E_\nu) - 2(E_{b_\ell} + E_\ell)\vec{e}_{b_\ell}^{\,\mathrm{T}} \cdot \vec{e}_\nu^{\,\mathrm{T}} - \\
&\quad 2p_\ell^z e_{b_\ell}^z - 2\vec{p}_\nu \cdot \vec{e}_{b_\ell} + 2\vec{p}_{b_\ell}^{\,\mathrm{T}} \cdot \vec{e}_{b_\ell}^{\,\mathrm{T}} \ , && \text{(D.44)} \\
\frac{\partial p_{b_\ell\nu}^z}{\partial E_{b_\ell}} &= e_{b_\ell}^z \ , && \text{(D.45)} \\
\frac{\partial m_{t_\ell}^2}{\partial p_\nu^z} &= \frac{\partial}{\partial p_\nu^z} \left(2p_\ell \cdot p_\nu + 2p_\ell \cdot p_{b_\ell} + 2p_\nu \cdot p_{b_\ell}\right) \\
&= 2\left(\frac{p_\nu^z}{E_\nu}\left(E_\ell + E_{b_\ell}\right) - p_\ell^z - p_{b_\ell^z}\right) \ , && \text{(D.46)} \\
\frac{\partial p_{b_\ell\nu}^z}{\partial p_\nu^z} &= 1 \ . && \text{(D.47)}
\end{aligned}$$

Concerning the generation of phase space points, a couple of integration variables are chosen to be fixed in order to speed up this relatively complex computation. These variables are the charged lepton energy E_ℓ, so that the finiteness of the lepton energy resolution is neglected, and the heavy particle masses squared, $m_{W_q}^2$, $m_{t_\ell}^2$ and $m_{t_q}^2$, where the narrow-width approximation is applied. The corresponding propagators which appear in the scattering amplitude must be adjusted accordingly,

$$\frac{1}{(m^2 - M^2)^2 + \Gamma^2 M^2} \xrightarrow[\Gamma/M \to 0]{} \frac{\pi}{\Gamma M} \delta(m^2 - M^2) \ . \qquad \text{(D.48)}$$

M denotes the propagator pole mass, Γ denotes the particle width and m denotes the particle mass. Altogether, the fixed variables are

$$\begin{aligned} E_\ell &= E_\ell^{\text{rec}} \ , \\ m_{W_q}^2 &= M_W^2 \ , \\ m_{t_\ell}^2 &= M_t^2 \ , \\ m_{t_q}^2 &= M_t^2 \ . \end{aligned} \tag{D.49}$$

There are five variables left which need to be computed from random numbers. These are the angles of the two unmatched final state light quarks and $p_{b_\ell\nu}^z$,

$$\begin{aligned} \cos\vartheta_{q/\,\bar{q}} &= 2\, r_{\cos\vartheta_{q/\,\bar{q}}} - 1 \ , \\ \varphi_{q/\,\bar{q}} &= 2\pi\, r_{\varphi_{q/\,\bar{q}}} \ , \\ p_{b_\ell\nu}^z &= \frac{\sqrt{s}}{3}\left(r_{p_{b_\ell\nu}^z} - \frac{1}{2}\right) \ . \end{aligned} \tag{D.50}$$

This way of generating $p_{b_\ell\nu}^z$ does not cover all possible configurations but restricts the phase space to the most important contributions. Combining these results, the phase space measure including the narrow-width approximations according to Eq. D.48 is

$$\begin{aligned} \int \mathrm{d}\Phi_{3,3}^{t\bar{t}_{\text{SL}}} &= \int \mathrm{d}x_1 \mathrm{d}x_2\, (2\pi)^4\, \delta^{(4)}\left(x_1 P_1 + x_2 P_2 - P\right) \\ &\quad \cdot \prod_{i\,\in\,\{\ell,\, b_\ell,\, b_q\}} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i} \delta\left(\cos\vartheta_i - \cos\vartheta_i^{\text{rec}}\right) \delta\left(\varphi_i - \varphi_i^{\text{rec}}\right) \delta(E_\ell - E_\ell^{\text{rec}}) \\ &\quad \cdot \prod_{j\,\in\,\{q,\, \bar{q},\, \nu\}} \frac{\mathrm{d}^3 p_j}{(2\pi)^3 2E_j} \left(\frac{\pi}{\Gamma_t M_t}\right)^2 \delta(m_{t_\ell}^2 - M_t^2) \delta(m_{t_q}^2 - M_t^2) \\ &\quad \cdot \frac{\pi}{\Gamma_W M_W}\, \delta(m_{W_q}^2 - M_W^2) \\ &= \frac{\pi^3}{32(2\pi)^{14}} \frac{1}{\Gamma_W M_W \Gamma_t^2 M_t^2} \frac{1}{s} E_\ell^{\text{rec}} \int \mathrm{d}E_q \mathrm{d}E_{\bar{q}} \mathrm{d}E_{b_\ell} \mathrm{d}E_{b_q} \mathrm{d}p_\nu^z \\ &\quad \cdot \mathrm{d}\cos\vartheta_q \mathrm{d}\varphi_q \mathrm{d}\cos\vartheta_{\bar{q}} \mathrm{d}\varphi_{\bar{q}} \frac{E_q E_{\bar{q}} E_{b_\ell} E_{b_q}}{E_\nu}\, \chi(0 < x_{1,2} < 1) \\ &\quad \cdot \delta(m_{t_\ell}^2 - M_t^2) \delta(m_{t_q}^2 - M_t^2) \delta(m_{W_q}^2 - M_W^2) \\ &= \frac{\pi^3}{24(2\pi)^{12}} \frac{E_\ell^{\text{rec}}}{\Gamma_W M_W \Gamma_t^2 M_t^2 \sqrt{s}} \int \mathrm{d}^5 r \left| \frac{\partial(m_{t_\ell}^2,\, m_{t_q}^2,\, m_{W_q}^2,\, p_{b_\ell\nu}^z)}{\partial(E_{\bar{q}},\, E_{b_\ell},\, E_{b_q},\, p_\nu^z)} \right|^{-1} \end{aligned}$$

$$\cdot \frac{E_q E_{\bar{q}} E_{b_\ell} E_{b_q}}{E_\nu} \chi(0 < x_{1,2} < 1)\chi(E_{b_q} > 0)\chi(E_{b_\ell\nu} > 0)\chi(p^2_{b_\ell\nu} > 0)\chi(E_{b_\ell} > 0) \ . \tag{D.51}$$

The characteristic functions at the end of this equation are needed since the mapping of random numbers given by Eq. D.50 does not necessarily lead to proper momentum configurations for any set of random numbers $r \in [0,\, 1]$.

References

[1] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, *Phys.Lett.*, B716:1–29, 2012.

[2] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, *Phys.Lett.*, B716:30–61, 2012.

[3] ATLAS Collaboration, ATLAS: Detector and physics performance technical design report. Volume 1, 1999, `https://cds.cern.ch/record/391176`.

[4] G. Bayatian et al., CMS Physics: Technical Design Report Volume 1: Detector Performance and Software, *CERN-LHCC-2006-001, CMS-TDR-8-1*, 2006, `https://cds.cern.ch/record/922757`.

[5] ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC, *CERN-LHCC-95-71, LHCC-P-3*, 1995, `http://cds.cern.ch/record/293391`.

[6] LHCb: Technical Proposal, *CERN-LHCC-98-004, LHCC-P-4*, 1998, `https://cds.cern.ch/record/622031`.

[7] The ATLAS, CDF, CMS and D0 Collaborations, First combination of Tevatron and LHC measurements of the top-quark mass, 2014, `http://arxiv.org/abs/1403.4427`.

[8] M. Veltman, Perturbation theory of massive Yang-Mills fields, *Nucl.Phys.*, B7: 637–650, 1968.

[9] G. 't Hooft, Renormalizable Lagrangians for Massive Yang-Mills Fields, *Nucl.Phys.*, B35:167–188, 1971.

[10] G. 't Hooft and M. Veltman, Regularization and Renormalization of Gauge Fields, *Nucl.Phys.*, B44:189–213, 1972.

[11] G. 't Hooft and M. Veltman, Combinatorics of gauge fields, *Nucl.Phys.*, B50: 318–353, 1972.

[12] C.-N. Yang and R. L. Mills, Conservation of Isotopic Spin and Isotopic Gauge Invariance, *Phys.Rev.*, 96:191–195, 1954.

[13] M. E. Peskin and D. V. Schroeder, An Introduction to quantum field theory, *Westview Press*, 1995.

[14] S. Glashow, Partial Symmetries of Weak Interactions, *Nucl.Phys.*, 22:579–588, 1961.

[15] S. Weinberg, A Model of Leptons, *Phys.Rev.Lett.*, 19:1264–1266, 1967.

[16] A. Salam, Elementary particle physics: Relativistic groups and analyticity, *Proceedings of the 8.th Nobel Symposium, Svartholm, N. (ed.), Stockholm*, 1968.

[17] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, *Phys.Rev.Lett.*, 13:321–323, 1964.

[18] P. W. Higgs, Broken symmetries, massless particles and gauge fields, *Phys.Lett.*, 12:132–133, 1964.

[19] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, *Phys.Rev.Lett.*, 13:508–509, 1964.

[20] G. Guralnik, C. Hagen, and T. Kibble, Global Conservation Laws and Massless Particles, *Phys.Rev.Lett.*, 13:585–587, 1964.

[21] P. W. Higgs, Spontaneous Symmetry Breakdown without Massless Bosons, *Phys.Rev.*, 145:1156–1163, 1966.

[22] T. Kibble, Symmetry breaking in nonAbelian gauge theories, *Phys.Rev.*, 155: 1554–1561, 1967.

[23] L. Alvarez-Gaume and J. Ellis, Eyes on a prize particle, *Nat Phys*, 7(1):2–3, Jan 2011, ISSN 1745-2473, URL `http://dx.doi.org/10.1038/nphys1874`.

[24] N. Cabibbo, Unitary Symmetry and Leptonic Decays, *Phys.Rev.Lett.*, 10:531–533, 1963.

[25] M. Kobayashi and T. Maskawa, CP Violation in the Renormalizable Theory of Weak Interaction, *Prog.Theor.Phys.*, 49:652–657, 1973.

[26] D. J. Gross and F. Wilczek, Ultraviolet Behavior of Nonabelian Gauge Theories, *Phys.Rev.Lett.*, 30:1343–1346, 1973.

[27] H. D. Politzer, Reliable Perturbative Results for Strong Interactions?, *Phys.Rev.Lett.*, 30:1346–1349, 1973.

[28] V. Gribov and L. Lipatov, Deep inelastic e p scattering in perturbation theory, *Sov.J.Nucl.Phys.*, 15:438–450, 1972.

[29] G. Altarelli and G. Parisi, Asymptotic Freedom in Parton Language, *Nucl.Phys.*, B126:298, 1977.

[30] Y. L. Dokshitzer, Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics., *Sov.Phys.JETP*, 46:641–653, 1977.

[31] A. Martin, W. Stirling, R. Thorne, and G. Watt, Parton distributions for the LHC, *Eur.Phys.J.*, C63:189–285, 2009.

[32] G. Watt, MSTW PDFs : key plots, website, `https://mstwpdf.hepforge.org/plots/plots.html`; as of April 27th 2015.

[33] T. Kinoshita, Mass singularities of Feynman amplitudes, *J.Math.Phys.*, 3:650–677, 1962.

[34] T. Lee and M. Nauenberg, Degenerate Systems and Mass Singularities, *Phys.Rev.*, 133:B1549–B1562, 1964.

[35] M. Cacciari, G. P. Salam, and G. Soyez, The Anti-k(t) jet clustering algorithm, *JHEP*, 0804:063, 2008.

[36] W. Stirling, private communication.

[37] CDF Collaboration, Observation of top quark production in $\bar{p}p$ collisions, *Phys.Rev.Lett.*, 74:2626–2631, 1995.

[38] D0 Collaboration, Observation of the top quark, *Phys.Rev.Lett.*, 74, 1995.

[39] D. Schaile and P. Zerwas, Measuring the weak isospin of B quarks, *Phys.Rev.*, D45: 3262–3265, 1992.

[40] H. Albrecht et al., Observation of B0 - anti-B0 Mixing, *Phys.Lett.*, B192:245–252, 1987.

[41] S. Glashow, J. Iliopoulos, and L. Maiani, Weak Interactions with Lepton-Hadron Symmetry, *Phys.Rev.*, D2:1285–1292, 1970.

[42] The LEP Collaborations: ALEPH, DELPHI, L3, OPAL, the LEP Electroweak Working Group, the SLD Electroweak Group, SLD Heavy Flavour Group, A Combination of preliminary electroweak measurements and constraints on the standard model, 2004, `http://arxiv.org/abs/hep-ex/0412015`.

[43] W. Bernreuther, Top quark physics at the LHC, *J.Phys.*, G35:083001, 2008.

[44] K. Olive et al., Review of Particle Physics, *Chin.Phys.*, C38:090001, 2014.

[45] J. Charles et al., CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories, *Eur.Phys.J.*, C41:1–131, 2005.

[46] ATLAS Collaboration, Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV, *Eur. Phys. J.*, C71:1577, 2011.

[47] CMS Collaboration, First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV, *Phys. Lett.*, B695:424–443, 2011.

[48] M. Czakon, P. Fiedler, and A. Mitov, Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$, *Phys.Rev.Lett.*, 110:252004, 2013.

[49] M. Czakon and A. Mitov, NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction, *JHEP*, 1301:080, 2013.

[50] M. Czakon and A. Mitov, NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels, *JHEP*, 1212:054, 2012.

[51] P. Baernreuther, M. Czakon, and A. Mitov, Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$, *Phys.Rev.Lett.*, 109:132001, 2012.

[52] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation, *Phys.Lett.*, B710:612–622, 2012.

[53] M. Czakon and A. Mitov, Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders, 2011, `http://arxiv.org/abs/1112.5675`.

[54] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, Hadronic top-quark pair production with NNLL threshold resummation, *Nucl.Phys.*, B855:695–741, 2012.

[55] A. Martin, W. Stirling, R. Thorne, and G. Watt, Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections, *Eur.Phys.J.*, C64:653–680, 2009.

[56] H.-L. Lai, M. Guzzi, J. Huston, et al., New parton distributions for collider physics, *Phys.Rev.*, D82:074024, 2010.

[57] J. Gao, M. Guzzi, J. Huston, et al., CT10 next-to-next-to-leading order global analysis of QCD, *Phys.Rev.*, D89(3):033009, 2014.

[58] R. D. Ball, V. Bertone, S. Carrazza, et al., Parton distributions with LHC data, *Nucl.Phys.*, B867:244–289, 2013.

[59] M. Botje, J. Butterworth, A. Cooper-Sarkar, et al., The PDF4LHC Working Group Interim Recommendations, 2011, `http://arxiv.org/abs/1101.0538`.

[60] LHC Top Working Group, Summary of LHC and Tevatron measurements of the top-quark pair production cross section, as of October 6th 2015, `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TopLHCWGSummaryPlots?rev=6`.

[61] C. D. White, S. Frixione, E. Laenen, and F. Maltoni, Isolating Wt production at the LHC, *JHEP*, 0911:074, 2009.

[62] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders, *Phys.Rev.Lett.*, 102:182003, 2009.

[63] J. M. Campbell and F. Tramontano, Next-to-leading order corrections to Wt production and decay, *Nucl.Phys.*, B726:109–130, 2005.

[64] J. M. Campbell, R. K. Ellis, and F. Tramontano, Single top production and decay at next-to-leading order, *Phys.Rev.*, D70:094012, 2004.

[65] B. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, The Fully differential single top quark cross-section in next to leading order QCD, *Phys.Rev.*, D66:054024, 2002.

[66] G. P. Lepage, A New Algorithm for Adaptive Multidimensional Integration, *J.Comput.Phys.*, 27:192, 1978.

[67] G. P. Lepage, VEGAS: AN ADAPTIVE MULTIDIMENSIONAL INTEGRATION PROGRAM, Cornell preprint CLNS-80/447, 1980.

[68] M. Lüscher, A Portable high quality random number generator for lattice field theory simulations, *Comput.Phys.Commun.*, 79:100–110, 1994.

[69] van Oldenborgh, Ff: A package to evaluate one loop feynman diagrams, *Comput.Phys.Commun 66*, 1991.

[70] P. Kant, O.M.Kind, T.Kintscher, T.Lohse, T.Martini, S.Mölbitz, P.Rieck, P.Uwer, HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions, *Comput.Phys.Commun.*, 191:74–89, 2015.

[71] N. Kidonakis, Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-, *Phys.Rev.*, D82:054018, 2010.

[72] LHC Top Working Group, Single top-quark reference cross sections, as of 2015, `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec?rev=19`.

[73] D0 Collaboration, Observation of Single Top Quark Production, *Phys.Rev.Lett.*, 103:092001, 2009.

[74] CDF collaboration, Observation of Single Top Quark Production and Measurement of —Vtb— with CDF, *Phys.Rev.*, D82:112005, 2010.

[75] The CDF and D0 Collaborations, Observation of s-channel production of single top quarks at the Tevatron, *Phys.Rev.Lett.*, 112:231803, 2014.

[76] T. M. Tait and C.-P. Yuan, Single top quark production as a window to physics beyond the standard model, *Phys.Rev.*, D63:014018, 2000.

[77] D0 Collaboration, Precision measurement of the ratio B($t \to Wb$)/B($t \to Wq$) and Extraction of V_{tb}, *Phys.Rev.Lett.*, 107:121802, 2011.

[78] CMS Collaboration, Measurement of the ratio $B(t \to Wb)/B(t \to Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV, *Phys.Lett.*, B736:33–57, 2014.

[79] H. Lacker, A. Menzel, F. Spettel, et al., Model-independent extraction of $|V_{tq}|$ matrix elements from top-quark measurements at hadron colliders, *Eur.Phys.J.*, C72:2048, 2012.

[80] The CDF and D0 collaborations, Tevatron Measurement of Single Top Quark Cross Sections and the CKM Matrix Element V_{tb}, 2014, `http://lss.fnal.gov/archive/2014/conf/fermilab-conf-14-370-e.pdf`.

[81] ATLAS Collaboration, Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Phys.Rev.*, D90(11):112006, 2014.

[82] CMS Collaboration, Measurement of the t-channel single-top-quark production cross section and of the $|\ V_{tb}\ |$ CKM matrix element in pp collisions at $\sqrt{s}$= 8 TeV, *JHEP*, 1406:090, 2014.

[83] M. C. Smith and S. Willenbrock, QCD and Yukawa corrections to single top quark production via $q\bar{q} \to t\bar{b}$, *Phys.Rev.*, D54:6696–6702, 1996.

[84] Q.-H. Cao, R. Schwienhorst, and C.-P. Yuan, Next-to-leading order corrections to single top quark production and decay at Tevatron. 1. s^- channel process, *Phys.Rev.*, D71:054023, 2005.

[85] G. Bordes and B. van Eijk, Calculating QCD corrections to single top production in hadronic interactions, *Nucl.Phys.*, B435:23–58, 1995.

[86] T. Stelzer, Z. Sullivan, and S. Willenbrock, Single top quark production via W - gluon fusion at next-to-leading order, *Phys.Rev.*, D56:5919–5927, 1997.

[87] T. Stelzer, Z. Sullivan, and S. Willenbrock, Single top quark production at hadron colliders, *Phys.Rev.*, D58:094021, 1998.

[88] Z. Sullivan, Understanding single-top-quark production and jets at hadron colliders, *Phys.Rev.*, D70:114012, 2004.

[89] Z. Sullivan, Angular correlations in single-top-quark and Wjj production at next-to-leading order, *Phys.Rev.*, D72:094034, 2005.

[90] Q.-H. Cao and C.-P. Yuan, Single top quark production and decay at next-to-leading order in hadron collision, *Phys.Rev.*, D71:054022, 2005.

[91] Q.-H. Cao, R. Schwienhorst, J. A. Benitez, R. Brock, and C.-P. Yuan, Next-to-leading order corrections to single top quark production and decay at the Tevatron: 2. t^- channel process, *Phys.Rev.*, D72:094027, 2005.

[92] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, NLO predictions for t-channel production of single top and fourth generation quarks at hadron colliders, *JHEP*, 0910:042, 2009.

[93] W. T. Giele, S. Keller, and E. Laenen, QCD corrections to W boson plus heavy quark production at the Tevatron, *Phys.Lett.*, B372:141–149, 1996.

[94] S. Zhu, Next-to-leading order QCD corrections to b g $\to$ t W^- at the CERN Large Hadron Collider, *Phys.Lett.*, B524:283–288, 2002.

[95] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, Single-top production in MC@NLO, *JHEP*, 0603:092, 2006.

[96] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, Single-top hadroproduction in association with a W boson, *JHEP*, 0807:029, 2008.

[97] S. Alioli, P. Nason, C. Oleari, and E. Re, NLO single-top production matched with shower in POWHEG: s- and t-channel contributions, *JHEP*, 0909:111, 2009.

[98] E. Re, Single-top Wt-channel production matched with parton showers using the POWHEG method, *Eur.Phys.J.*, C71:1547, 2011.

[99] S. Mrenna and C. Yuan, Effects of QCD resummation on W+ h and t anti-b production at the Tevatron, *Phys.Lett.*, B416:200–207, 1998.

[100] N. Kidonakis, Single top production at the Tevatron: Threshold resummation and finite-order soft gluon corrections, *Phys.Rev.*, D74:114012, 2006.

[101] N. Kidonakis, Higher-order soft gluon corrections in single top quark production at the LHC, *Phys.Rev.*, D75:071501, 2007.

[102] N. Kidonakis, Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution, *Phys.Rev.*, D82:114030, 2010.

[103] N. Kidonakis, Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production, *Phys.Rev.*, D83:091503, 2011.

[104] T. Kintscher, Fast single-top cross-section predictions for hadron colliders with the hathor program, *M.Sc. thesis, Humboldt-Universität zu Berlin*, 2013.

[105] M. Brucherseifer, F. Caola, and K. Melnikov, On the NNLO QCD corrections to single-top production at the LHC, *Phys.Lett.*, B736:58–63, 2014.

[106] M. Assadsolimani, P. Kant, B. Tausk, and P. Uwer, Calculation of two-loop QCD corrections for hadronic single top-quark production in the t channel, *Phys.Rev.*, D90(11):114024, 2014.

[107] P. Kant et al., HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions, *e-print arXiv:1406.4403*, 2015, `http://arxiv.org/abs/1406.4403`.

[108] M. Aliev, H. Lacker, U. Langenfeld, et al., HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR, *Comput.Phys.Commun.*, 182:1034–1046, 2011.

[109] S. Alekhin, J. Blumlein, and S. Moch, Parton Distribution Functions and Benchmark Cross Sections at NNLO, *Phys.Rev.*, D86:054009, 2012.

[110] R. D. Ball, L. Del Debbio, S. Forte, et al., A first unbiased global NLO determination of parton distributions and their uncertainties, *Nucl.Phys.*, B838:136–206, 2010.

[111] A. Buckley, J. Butterworth, S. Gieseke, et al., General-purpose event generators for LHC physics, *Phys.Rept.*, 504:145–233, 2011.

[112] CDF collaboration, Search for s-Channel Single-Top-Quark Production in Events with Missing Energy Plus Jets in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.96 TeV, *Phys.Rev.Lett.*, 112(23):231805, 2014.

[113] D0 Collaboration, Evidence for s-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys.Lett.*, B726:656–664, 2013.

[114] The CDF and D0 collaborations, Observation of s-channel production of single top quarks at the tevatron, *website*, `http://tevewwg.fnal.gov/singleTop/public/schannel/`, as of May 15th 2015.

[115] CMS Collaboration, Search for s-channel single top-quark production in pp collisions at sqrt(s) = 8 TeV, *CMS-PAS-TOP-13-009*, 2013, `https://cds.cern.ch/record/1633190`.

[116] ATLAS Collaboration, Search for s-channel single top-quark production in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, *Phys.Lett.*, B740: 118–136, 2015.

[117] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, *JHEP*, 1006:043, 2010.

[118] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with Parton Shower simulations: the POWHEG method, *JHEP*, 0711:070, 2007.

[119] P. Nason, A New method for combining NLO QCD with shower Monte Carlo algorithms, *JHEP*, 0411:040, 2004.

[120] S. Frixione and B. R. Webber, Matching NLO QCD computations and parton shower simulations, *JHEP*, 0206:029, 2002.

[121] T. Sjostrand and M. van Zijl, A Multiple Interaction Model for the Event Structure in Hadron Collisions, *Phys.Rev.*, D36:2019, 1987.

[122] H. Schulz, *Measurement of the Underlying Event using track-based event shapes in* $Z \to \ell^+\ell^-$ *events with ATLAS*, PhD thesis, Humboldt-Universität zu Berlin, 2015.

[123] ATLAS Collaboration, Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC, *New J.Phys.*, 13:053033, 2011.

[124] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Parton Fragmentation and String Dynamics, *Phys.Rept.*, 97:31–145, 1983.

[125] B. Andersson, The Lund model, *Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol.*, 7: 1–471, 1997.

[126] D. Amati and G. Veneziano, Preconfinement as a Property of Perturbative QCD, *Phys.Lett.*, B83:87, 1979.

[127] J. Alwall, R. Frederix, S. Frixione, et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *JHEP*, 1407:079, 2014.

[128] B. P. Kersevan and E. Richter-Was, The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1, *Comput.Phys.Commun.*, 184:919–985, 2013.

[129] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, ALPGEN, a generator for hard multiparton processes in hadronic collisions, *JHEP*, 0307:001, 2003.

[130] G. Corcella, I. Knowles, G. Marchesini, et al., HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes), *JHEP*, 0101:010, 2001.

[131] M. Bahr, S. Gieseke, M. Gigg, et al., Herwig++ Physics and Manual, *Eur.Phys.J.*, C58:639–707, 2008.

[132] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 Physics and Manual, *JHEP*, 0605:026, 2006.

[133] T. Sjostrand, S. Mrenna, and P. Z. Skands, A Brief Introduction to PYTHIA 8.1, *Comput.Phys.Commun.*, 178:852–867, 2008.

[134] T. Gleisberg, S. Hoeche, F. Krauss, et al., Event generation with SHERPA 1.1, *JHEP*, 0902:007, 2009.

[135] LEP design report, *CERN-LEP-84-01*, 1984, `https://cds.cern.ch/record/102083`.

[136] O. S. Brüning, P. Collier, P. Lebrun, et al., LHC Design Report, *CERN-2004-003-V-1*, 2004, `https://cds.cern.ch/record/782076`.

[137] L. Evans and P. Bryant, LHC Machine, *JINST*, 3:S08001, 2008.

[138] F. Marcastel, CERN's Accelerator Complex., 2013, `https://cds.cern.ch/record/1621583`, modified version.

[139] ATLAS Collaboration, website, `http://www.atlas.ch`.

[140] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, *JINST*, 3:S08003, 2008.

[141] D. Green (ed.), At the leading edge: The ATLAS and CMS LHC experiments, 2010.

[142] ATLAS Collaboration, ATLAS liquid-argon calorimeter: Technical Design Report, *ATLAS-TDR-2, CERN-LHCC-96-041*, 1996, `https://cds.cern.ch/record/331061`.

[143] ATLAS Collaboration, Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV, *Eur. Phys. J.*, C75(3):120, 2015.

[144] ATLAS Collaboration, ATLAS muon spectrometer: Technical Design Report, *ATLAS-TDR-10, CERN-LHCC-97-022*, 1997, https://cds.cern.ch/record/331068.

[145] Worldwide LHC Computing Grid, website, http://wlcg.web.cern.ch/.

[146] ATLAS Collaboration, Improved Luminosity Determination in **pp** Collisions at $\sqrt{\mathbf{s}} = \mathbf{7}$ TeV using the ATLAS Detector at the LHC, *ATLAS-CONF-2012-080*, 2012, http://cds.cern.ch/record/1460392.

[147] ATLAS Collaboration, The ATLAS Simulation Infrastructure, *Eur.Phys.J.*, C70: 823–874, 2010.

[148] S. Agostinelli et al., GEANT4: A Simulation toolkit, *Nucl.Instrum.Meth.*, A506: 250–303, 2003.

[149] ATLAS Collaboration, Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment, *ATLAS-CONF-2012-042*, 2012, https://cds.cern.ch/record/1435196?ln=en.

[150] ATLAS Collaboration, Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment, *ATLAS-CONF-2010-069*, 2010, http://cds.cern.ch/record/1281344.

[151] ATLAS Collaboration, Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data, *ATLAS-CONF-2014-032*, 2014, http://cds.cern.ch/record/1706245.

[152] ATLAS Collaboration, Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung, (ATLAS-CONF-2012-047), 2012, http://cds.cern.ch/record/1449796.

[153] ATLAS Collaboration, Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data, *Eur.Phys.J.*, C74(10):3071, 2014.

[154] ATLAS Collaboration, Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data, *Eur.Phys.J.*, C74(11):3130, 2014.

[155] ATLAS Collaboration, Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Eur.Phys.J.*, C75(1):17, 2015.

[156] ATLAS Collaboration, Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV, *Eur.Phys.J.*, C73(3):2304, 2013.

[157] ATLAS Collaboration, Pile-up subtraction and suppression for jets in ATLAS, *ATLAS-CONF-2013-083*, 2013, `http://cds.cern.ch/record/1570994`.

[158] P. Bagnaia et al., Measurement of Jet Production Properties at the CERN anti-p p Collider, *Phys.Lett.*, B144:283, 1984.

[159] ATLAS Collaboration, Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector, *Eur.Phys.J.*, C73(3):2306, 2013.

[160] A. Hoecker, P. Speckmayer, J. Stelzer, et al., TMVA: Toolkit for Multivariate Data Analysis, *PoS*, ACAT:040, 2007.

[161] ATLAS Collaboration, Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment, *ATLAS-CONF-2014-004*, 2014, `https://cds.cern.ch/record/1664335`.

[162] ATLAS Collaboration, Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data, 2014, `http://cds.cern.ch/record/1741020`.

[163] ATLAS Collaboration, Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV, *ATLAS-CONF-2013-082*, 2013, `https://cds.cern.ch/record/1570993`.

[164] E. Barberio and Z. Was, PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0, *Comput.Phys.Commun.*, 79:291–308, 1994.

[165] S. Jadach, Z. Was, R. Decker, and J. H. Kuhn, The tau decay library TAUOLA: Version 2.4, *Comput.Phys.Commun.*, 76:361–380, 1993.

[166] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO, *Phys.Rev.*, D69:094008, 2004.

[167] J. M. Campbell, R. K. Ellis, and C. Williams, Vector boson pair production at the LHC, *JHEP*, 1107:018, 2011.

[168] ATLAS Collaboration, Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s}$=8 TeV

with the ATLAS detector, *ATLAS-CONF-2014-058*, 2014, https://cds.cern.ch/record/1951336.

[169] C. Monini, A. Lleres, and A. Lucotte, *Single-top s channel cross-section measurement with the ATLAS detector*, PhD thesis, Joseph Fourier U., 2014.

[170] S. Stamm, private communication, Humboldt-Universität zu Berlin. PhD thesis in preparation.

[171] K. Kondo, Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models, *J. Phys. Soc. Jap.*, 57:4126–4140, 1988.

[172] K. Kondo, Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. 2: Mass spectra for $2 \to 2$ processes, *J. Phys. Soc. Jap.*, 60:836–844, 1991.

[173] D0 Collaboration, Precision measurement of the top-quark mass in lepton+jets final states, *Phys.Rev.*, D91(11):112003, 2015.

[174] CDF Collaboration, Measurement of w-boson polarization in top-quark decay using the full cdf run ii data set, *Phys. Rev. D*, 87:031104, Feb 2013.

[175] ATLAS Collaboration, Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, *e-print arXiv:1503.05066, CERN-PH-EP-2015-047*, 2015, http://cds.cern.ch/record/2001975.

[176] V. Khachatryan et al., Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method, *Eur.Phys.J.*, C75(6):251, 2015.

[177] J. Erdmann, S. Guindon, K. Kroeninger, et al., A Likelihood-Based Reconstruction Algorithm for Top-Quark Pairs and the KLFitter Framework, *Nucl. Instrum. Meth.*, A748:18–25, 2014.

[178] J. Neyman and E. Pearson, On the Problem of the Most Efficient Tests of Statistical Hypotheses, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 1933, ISSN 02643952, URL http://www.jstor.org/stable/91247.

[179] K. Cranmer and T. Plehn, Maximum Significance at the LHC and Higgs Decays to Muons, *Eur. Phys. J.*, C51:415–420, 2007.

[180] T. Martini and P. Uwer, Extending the Matrix Element Method beyond the Born approximation: Calculating event weights at next-to-leading order accuracy, *JHEP*, 09:083, 2015.

[181] P. Artoisenet and O. Mattelaer, MadWeight: Automatic event reweighting with matrix elements, *PoS*, CHARGED2008:025, 2008.

[182] P. Artoisenet, P. de Aquino, F. Maltoni, and O. Mattelaer, Unravelling $t\bar{t}h$ via the Matrix Element Method, *Phys.Rev.Lett.*, 111(9):091802, 2013.

[183] P. Artoisenet, V. Lemaitre, F. Maltoni, and O. Mattelaer, Automation of the matrix element reweighting method, *JHEP*, 1012:068, 2010.

[184] Rieck, P. et. al., The mem for analyses at the lhc, *Presentation at the Zurich Phenomenology Workshop, University of Zurich*, 2014, `https://indico.cern.ch/event/280658/contribution/15`, as of October 10th 2015.

[185] T. Hahn, The CUBA Library, *Nucl. Instrum. Meth.*, A559:273–277, 2006.

[186] M. Whalley, D. Bourilkov, and R. Group, The Les Houches accord PDFs (LHAPDF) and LHAGLUE, *e-print*, 2005, `http://arxiv.org/abs/hep-ph/0508110`.

[187] J. M. Campbell and R. Ellis, MCFM for the Tevatron and the LHC, *Nucl.Phys.Proc.Suppl.*, 205-206:10–15, 2010.

[188] Root data analysis framework, *website*, `http://root.cern.ch/drupal`, as of July 1st 2015.

[189] A. Caldwell, D. Kollar, and K. Kröninger, BAT: The Bayesian Analysis Toolkit, *Comput. Phys. Commun.*, 180:2197–2209, 2009.

[190] ATLAS Collaboration, Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector using the Matrix Element Method, *ATLAS-CONF-2015-047*, 2015, `http://cds.cern.ch/record/2052585`.

[191] ATLAS Collaboration, Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS, *ATLAS-CONF-2015-037*, 2015, `http://cds.cern.ch/record/2044941`.

[192] ATLAS Collaboration, Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy of $\sqrt{s} = 7$ TeV, *ATLAS-CONF-2010-054*, 2010, `https://cds.cern.ch/record/1281311`.

[193] S. Ellis, R. Kleiss, and W. J. Stirling, W's, Z's and Jets, *Phys. Lett.*, B154:435, 1985.

[194] F. A. Berends, H. Kuijf, B. Tausk, and W. Giele, On the Production of a W and Jets at Hadron Colliders, *Nucl. Phys.*, B357:32–64, 1991.

[195] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, HistFactory: A tool for creating statistical models for use with RooFit and RooStats, *CERN-OPEN-2012-016*, 2012, `https://cds.cern.ch/record/1456844`.

[196] L. Moneta, K. Belasco, K. S. Cranmer, et al., The RooStats Project, *PoS*, ACAT2010:057, 2010.

[197] W. Verkerke and D. P. Kirkby, The RooFit Toolkit for Data Modeling, *CHEP-2003-MOLT007*, C0303241, 2003.

[198] R. J. Barlow and C. Beeston, Fitting Using Finite Monte Carlo Samples, *Comput. Phys. Commun.*, 77:219–228, 1993.

[199] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic Formulae for Likelihood-Based Tests of New Physics, *Eur. Phys. J.*, C71:1554, 2011.

[200] S. Catani and M. Seymour, A General algorithm for calculating jet cross-sections in NLO QCD, *Nucl.Phys.*, B485:291–419, 1997.

[201] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, The Dipole formalism for next-to-leading order QCD calculations with massive partons, *Nucl.Phys.*, B627: 189–265, 2002.

[202] T. Martini, private communication, Humboldt-Universität zu Berlin.

[203] S. Mölbitz, Berechnung von korrekturen hherer ordnung in der qcd zur erzeugung einzelner top-quarks, *M.Sc. thesis, private communication, Humboldt-Universität zu Berlin*, 2013.

[204] ATLAS Collaboration, Measurement of the Inclusive and Fiducial Cross-Section of Single Top-Quark t-Channel Events in pp Collisions at $\sqrt{s} = 8$ TeV, *ATLAS-*

CONF-2014-007, 2014, http://cds.cern.ch/record/1668960.

[205] CMS Collaboration, Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV, *JHEP*, 1212:035, 2012.

[206] ATLAS Collaboration, Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV, *Phys.Lett.*, B716:142–159, 2012.

[207] ATLAS Collaboration, Measurement of the cross-section for associated production of a top quark and a W boson at $\sqrt{s} = 8$ TeV with the ATLAS detector, *ATLAS-CONF-2013-100*, 2013, http://cds.cern.ch/record/1600799.

[208] The ATLAS and CMS collaborations, Combination of cross-section measurements for associated production of a single top-quark and a W boson at $\sqrt{s} = 8$ TeV with the ATLAS and CMS experiments, (ATLAS-CONF-2014-052), Sep 2014, URL http://cds.cern.ch/record/1951032.

[209] CMS Collaboration, Evidence for associated production of a single top quark and W boson in pp collisions at $\sqrt{s} = 7$ TeV, *Phys.Rev.Lett.*, 110:022003, 2013.

[210] CMS Collaboration, Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}$ =8 TeV, *Phys.Rev.Lett.*, 112(23):231802, 2014.

[211] ATLAS Collaboration, Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector, *J. High Energy Phys.*, 01:069. 69 p, 2014, https://cds.cern.ch/record/1757231.

[212] P. Schieferdecker, *Measurement of the Top Quark Mass at D0 Run II with the Matrix Element Method in the Lepton+Jets Final State*, PhD thesis, Ludwig-Maximilians-Universität München, 2005.

[213] A. Buckley, The hepthesis LaTeX class.

List of Figures

1.1. Yukawa interaction vertex 7
1.2. Higgs Potential 12
1.3. Factorization theorem 19
1.4. MSTW2008 Proton PDF set. 21
1.5. Jet clustering 23
1.6. Cross sections at hadron colliders 25
1.7. Top-quark decay 27
1.8. Top pair production Feynman diagrams 28
1.9. $t\bar{t}$ event topology 29
1.10. $t\bar{t}$ cross sections 29
1.11. Single top-quark production Feynman diagrams 31
1.12. CKM matrix elements in single top-quark production 33
1.13. Single top-quark production cross sections, s-channel 36
1.14. Scale dependencies of the s-channel cross section 36
1.15. Top-quark production cross sections as a function of the top-quark mass 38
1.16. PDF uncertainties of s-channel single top-quark production cross sections 38
1.17. Overview of searches for s-channel single top-quark production 41
1.18. Collision event simulation 43

2.1. CERN's accelerator complex 51
2.2. ATLAS detector 55
2.3. ATLAS infrastructure 56
2.4. Illustration of the pseudorapidity 57
2.5. ATLAS inner detector, barrel 59
2.6. ATLAS inner detector, barrel and end-cap 61
2.7. ATLAS calorimeters 62
2.8. ATLAS electromagnetic calorimeter cells 64
2.9. ATLAS hadronic tile calorimeter module 65
2.10. ATLAS muon spectrometer 67
2.11. ATLAS muon spectrometer monitored drift tube module 68
2.12. ATLAS trigger scheme 71
2.13. ATLAS integrated luminosity in 2012 74
2.14. Number of interactions per bunch crossing in 2012 76
2.15. ATLAS electron reconstruction efficiency 79
2.16. ATLAS electron identification efficiency 81
2.17. ATLAS muon reconstruction efficiency 85
2.18. ATLAS muon momentum calibration 86
2.19. ATLAS muon trigger efficiency 87
2.20. ATLAS jet energy response before the final calibration 89
2.21. ATLAS absolute jet energy in-situ calibration 90
2.22. ATLAS jet energy scale uncertainty 92
2.23. Impact parameters of tracks 95
2.24. ATLAS b-tagging calibration. 96

3.1. Analysis infrastructure 100
3.2. Signal and background processes 103
3.3. ATLAS prompt lepton fake estimation 108
3.4. Definition of the W+jets validation region 113
3.5. Selected event display 114
3.6. Kinematic distributions in the signal region 118
3.7. Kinematic distributions in the W+jets region 119
3.8. Kinematic distributions in the $t\bar{t}$ region 120
3.9. Resolution functions 140
3.10. Transfer function efficiencies 142
3.11. Transfer function efficiencies, continued 143
3.12. TELL control distributions, signal region 144
3.13. TELL control distributions, W+jets control region 145
3.14. TELL control distributions, $t\bar{t}$ control region 146
3.15. Expected signal discriminant distribution 148
3.16. Final expected signal discriminant distribution 150
3.17. Signal probability control distributions 150
3.18. Impact of transfer functions on the TELL discriminant 153
3.19. Approximation of signal probabilities 155
3.20. TELL discriminant shape variation examples 164
3.21. Systematic shape variation p-values 166
3.22. Profile likelihood curve 171
3.23. Nuisance parameter fit results 173
3.24. Fit parameter correlations 174
3.25. Post-fit discriminant distribution 175

3.26. Post-fit discriminant distribution with background subtraction 175
3.27. Search for s-channel single top-quark production 177
3.28. Signal cross section measurement 179
3.29. Cross sections relevant to s-channel measurements in pp collisions. 185
3.30. Expected TELL discriminant distribution at LHC Run 2 185
3.31. Expected signal significance at LHC Run 2 187
3.32. Expected signal cross section uncertainty at LHC Run 2 187

B.1. Further single top-quark production cross sections 206
B.2. Further scale dependencies of single top-quark production cross sections . 207
B.3. PDF uncertainties of single top-quark production cross sections 208

C.1. Sub-leading jet control distributions 212
C.2. Electron control distributions 213
C.3. Muon control distributions 214
C.4. $t\bar{t}$ decay channels 215
C.5. Electron energy resolution functions 217
C.6. Muon p_{T} resolution functions 218
C.7. b-jet energy resolution functions 219
C.8. TELL discriminant binning 221
C.9. TELL discriminant validation in the signal region 222
C.10. Impact of discriminant shape variations on the signal significance. 224
C.11. Discriminant distribution predicted by different W+jets generators . . . 225
C.12. Fit of the discriminant using the Alpgen +Pythia 6 W+jets prediction 226
C.13. Comparison of fitted signal strengths using different W+jets predictions 226
C.14. Nuisance parameter pulls, extended 227

C.15.Nuisance parameter pulls, extended 228

C.16.Nuisance parameter pulls, extended 229

List of Tables

1.1. Electroweak quantum numbers of fermions 15

1.2. Selected top-quark cross sections at hadron colliders 39

1.3. Selected event generators 47

2.1. LHC design beam parameters 54

3.1. Event selection of signal and validation regions 112

3.2. Event yields in the signal region 116

3.3. Event yields in the W+jets control region 117

3.4. Event yields in the $t\bar{t}$ control region 117

3.5. Overview of signal separation methods 127

3.6. Considered TELLs 138

3.7. Performance of TELL computations 151

3.8. Rate uncertainties in the signal region 163

3.9. Discriminant shape variations due to systematic uncertainties 167

3.10. Break-down of systematic uncertainties 181

A.1. Nominal MC samples 195

A.2. MC samples of modelling studies 196

A.3. MC samples of PDF studies 197

A.4. MC samples used in transfer function studies 197

C.1. Control distribution fit results, signal region 210

C.2. Control distribution fit results, W+jets region 210

C.3. Control distribution fit results, $t\bar{t}$ region 211

C.4. Event yields and di-lepton veto in the signal region 216

Colophon

This thesis was made in LaTeX using the "hepthesis" class [213].

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig und nur unter Zuhilfenahme der von mir angegebenen Hilfsmittel erarbeitet und verfasst habe.

Berlin, 13.10.2015 Patrick Rieck

Danksagung

First of all I would like to thank the generation of physicists who designed and constructed the LHC and the ATLAS detector. It was a pleasure to work with this fascinating device!

Ich danke außerdem Prof. Thomas Lohse für seine anhaltende Unterstützung und auch dafür, dass er bei der Wahl unseres Forschungsprojekts den richtigen Riecher hatte. Auch meinen Kollegen Oliver Maria Kind und Sören Stamm gilt mein Dank für unsere produktive Zusammenarbeit. Danke auch an Sebastian Mergelmeyer für die Unterstützung beim Erstellen der Dissertation. Vielen Dank ebenso an die vielen Kollegen der ATLAS Kollaboration, die mich unterstützt, angeleitet und aufgemuntert haben. Es gibt eine Reihe von Institutionen, die meine Arbeit ermöglicht haben, wobei ich insbesondere der Studienstiftung danke, die mich außerdem auf meinem Weg begleitet hat.

Am Institut für Physik der HU gab und gibt es noch viele weitere Leute, die meine Arbeit ermöglicht oder zumindest erleichtert haben. Allen voran Veronika Schneider mit ihrer Professionalität sowie Freundlichkeit. Hinzu kommen hier viele weitere Mitstreiter, die ich einfach aufzählen möchte: Carsten Kendziorra, Dennis Wendland, Geoffrey Herbert, Holger Schulz, Laura Rehnisch, Lukas Heinrich, Thomas Kintscher, Rocco Mandrysch & Co. – vor allem unsere gemeinsamen Besuche von Konferenzen und Workshops haben mich persönlich weitergebracht. Gern denke ich auch an meine Zeit in Genf zurück – lang leben die dreisten drei!

Mein besonderer Dank gilt meinen Freundinnen und Freunden außerhalb der Physik und vor allem meiner Familie für ihre unbedingte Unterstützung und Geduld.

This thesis presents a measurement of s-channel single top-quark production in proton-proton collisions. The data set has been recorded by the ATLAS detector at the LHC in 2012. Due to the rareness of the signal collision events, the method of total event likelihoods is applied in order to separate the signal from large background contributions. By means of a statistical evaluation of the data, the presence of the signal process is observed with a significane of more than three standard deviations. This is the first evidence for s-channel single top-quark production in proton-proton collisions. The measurement agrees with the standard model prediction.

Support

Studienstiftung des deutschen Volkes

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

HELMHOLTZ GEMEINSCHAFT

DFG Deutsche Forschungsgemeinschaft